THE UNIVERSITY OF CHICAGO

SEARCHING FOR NEW PHYSICS USING DI-HIGGS FINAL STATES WITH THE ATLAS DETECTOR

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY
MICHAEL HANK

CHICAGO, ILLINOIS
DECEMBER 2022

To all those who have taught me, all those who have supported me, and all those who have shaped me on my journey through life.

"We all love a good story. We all love a tantalizing mystery. We all love the underdog pressing onward against seemingly insurmountable odds. We all, in one form or another, are trying to make sense of the world around us. And all of these elements lie at the core of modern physics. The story is among the grandest – the unfolding of the entire universe; the mystery is among the toughest – finding out how the cosmos came to be; the odds are among the most daunting – bipeds, newly arrived by cosmic time scales trying to reveal the secrets of the ages; and the quest is among the deepest – the search for fundamental laws to explain all we see and beyond, from the tiniest particles to the most distant galaxies."

– Brian Greene, The Elegant Universe: Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory

Table of Contents

IV Di-Higgs Analyses

V Conclusions

Appendix

List of Figures

List of Tables

ACKNOWLEDGMENTS

I would like to express my utmost thanks to my advisor, Professor Young-Kee Kim. Her guidance has been essential both in helping me complete this work as well as preparing me for the next steps in my future. I am grateful for all of the time and insight she has given me over the years. I would also like to thank Professor Melvyn Shochet and Professor Lauren Tompkins for their direction and oversight on my hardware projects. I also thank Professor David Schmitz, Professor Marcela Carena, and Professor Cheng Chin for providing feedback on this work.

Thanks are extended to all members of the multi-b and di-Higgs analysis teams, whose contributions made this work possible. In particular, I would like to thank Marco Valente and Carlos Moreno Martinez for all of their work bringing this analysis to fruition. Special thanks are extended to Maximilian Swiatlowski, who always went above and beyond in guiding me through these analyses and provided me with invaluable advice, support, and mentorship.

I would also like to thank my former and undergraduate advisors, Professor Luca Grandi and Professor Lucy Fortson, for sparking my interest in research and starting me on the path to where I am today.

I thank the numerous teachers and professors who have taught, encouraged, and inspired me throughout the years. Any list will be woefully incomplete, but Carlton Urdahl, Charley Keifenheim, Denise Wahlin-Fiskum, Kip Wold, Peggy Pasche, Steve Hermanson, and Vicki Cary merit special mention.

My parents, Janice and Patrick, have my deepest and heartfelt thanks. Their support provided the foundation making all of this possible.

To Louis: These past months have meant more to me than I can put into words. I love you.

ABSTRACT

Searches for new physics using di-Higgs final states are presented. These states are investigated in two contexts. The first analysis targets pairs of Higgs bosons decaying to $b\bar{b}b\bar{b}$, produced with missing energy from the decays of supersymmetric higgsinos. This follows up on an excess at 250 GeV observed using 24.3 fb^{-1} of data and searches for higgsinos in the well-motivated low-mass regime. The second analysis targets pairs of Higgs bosons produced via the trilinear Higgs self-coupling and considers multiple final states, with an emphasis on the $b\bar{b}b\bar{b}$ channel. The searches are performed using 126 fb^{-1} and up to 139 fb^{-1} of data respectively in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector. The $b\bar{b}b\bar{b}$ backgrounds are dominated by difficult-to-model QCD multijet events and are estimated using novel data-driven approaches utilizing machine-learning techniques. No significant excess above the Standard Model is observed. For pure higgsino to Higgs boson decays, the supersymmetry analysis excludes higgsino masses up to 530 GeV at a 95% confidence level. Interpretations in alternate scenarios as well model-independent limits are also presented. For the Higgs self-coupling analysis, the self-coupling parameter is constrained to $\lambda_{hhh} \in [-3.9\lambda_{hhh}^{\mathrm{SM}}, 11.1\lambda_{hhh}^{\mathrm{SM}}]$ at a 95% confidence level for the $b\bar{b}b\bar{b}$ channel and $\lambda_{hhh} \in [-0.4\lambda_{hhh}^{\mathrm{SM}}, 6.3\lambda_{hhh}^{\mathrm{SM}}]$ for the combined channels.

Part I

Introduction and Theory

Chapter 1

INTRODUCTION

For much of human history, philosophers and scientists have pondered a deceptively simple question: What are the smallest, most fundamental building blocks of our universe? The idea that there are some smallest such constituents dates back to at least ∼400 BCE, when the Greek philosopher Democritus proposed the universe is made up of 'atomos', which translates to 'indivisible' [1]. When John Dalton proposed his atomic theory in the early 1800s, arguing from a more scientific basis that all matter was built from a handful of indivisible building blocks, it was only natural that he borrowed this term. The idea of atoms as fundamental particles became established in the scientific community.

Nature, it turned out, went a level deeper.

In 1897, J. J. Thomson discovered the electron and realized it was a *sub*-atomic particle [2]. The atomic model thus had to be revised: instead of being indivisible, atoms were now described by a "plum pudding" model where they were divided into a uniform positive charge and embedded negative electrons. This model had to be revised again a decade later, when the Rutherford gold foil experiments showed via scattering that the positive charge was concentrated in a nucleus much smaller than the atom itself. The last revision to this era of particle physics came with the discovery of the neutron by James Chadwick in 1932. Atoms were then described by the Bohr model, with three elementary building blocks: protons and neutrons (forming the nuclei), and their orbiting electrons [3].

Nature, it turned out, went yet another level deeper.

In 1960s and 1970s, it was found that protons and neutrons had smaller constituents called "quarks", fractionally charged particles with half-integer spins. Protons were made of two "up" quarks (+2/3 charge) and one "down" quark (-1/3 charge), while neutrons were made of one up quark and two down quarks. Several elementary particles not present in typical atoms were also discovered in the intervening years or shortly thereafter. These included two charged leptons, the muon (μ) and tau (τ), which are identical to the electron except for their

heavier masses. These were complemented by neutral leptons of low mass called "neutrinos" (ν), with a type corresponding to each of the charged leptons. Four heavier quarks were also discovered, with the "charm" and "top" quarks as heavier counterparts to the up quark and the "strange" and "bottom" quarks as heavier counterparts to the down quark. In addition to these half-integer spin particles (called fermions), 4 bosons (integer-spin particles) were also discovered: the photon, the gluon, and the $W^{\pm}$ and Z bosons. Finally, each particle had a corresponding antiparticle (except the neutral bosons, which served as their own) of the same mass but opposite quantum numbers, such as charge [3].

These particles and their rules of interaction were organized into what has become known as the Standard Model (SM) of particle physics [3]. The SM consists of 3 generations of fermions, with each generation identical in all ways except mass. Each generation consists of a +2/3 and a -1/3 charged quark, as well as a -1 charged lepton and its corresponding neutrino. The interactions of these particles are governed by 3 forces, mediated by the bosons. The electromagetic force is carried by the photon and allows interactions between charged particles; this is the force that attracts electrons to nuclei. The strong nuclear force governs interactions between particles exhibiting "color charge", i.e. quarks and gluons. This force is mediated by gluons and is responsible for binding protons and neutrons into nuclei. Finally, the weak nuclear force, carried by the $W^{\pm}$ and Z bosons, allows interactions between any particle based on their weak hypercharge. This interaction is responsible for β-decay, and is the only force that can convert particles from one generation to another.

The Standard Model required the existence of one more elementary particle. In order to generate the observed masses of the $W^{\pm}$ and Z bosons, a spin-0 boson called the Higgs boson was necessary. After decades of searches, the Higgs boson was observed in 2012 by the ATLAS and CMS experiments at the Large Hadron Collider, providing the final linchpin to the Standard Model [4, 5]. This theory has been enormously successful, correctly describing nearly every physical process observed to astonishing precision.

However, there are signs that the Standard Model is incomplete. As examples (to be

discussed in more detail in Chapter 2), it does not account for gravity, the existence of dark matter, the smallness of the Higgs boson mass, nor the imbalance between matter and anti-matter. Various models have been proposed to account for these problems. Many of these theories, such as supersymmetry (SUSY), introduce new elementary particles. The goal of this thesis is to investigate possible Beyond the Standard Model (BSM) physics through final states involving pairs of Higgs bosons.

It may be that the universe has yet more building blocks, with us only having glimpsed a fraction of an overarching pattern. It could even be that, like the "indivisible" atoms, our elementary particles aren't elementary at all. Either way, it seems that nature will go yet another level deeper.

This thesis is organized into 5 parts. Part I introduces the thesis and (Chapter 2) provides theoretical background and motivation. Part II describes the experimental setup. Chapter 3 discusses the Large Hadron Collider and Chapter 4 describes the ATLAS detector. Part III describes the $HH \to b\bar{b}b\bar{b} + E_{\mathrm{T}}^{\mathrm{miss}}$ search targeting supersymmetric di-Higgs production. Chapter 5 discusses the ATLAS data and generated signal models. Chapter 6 defines the objects used in the analysis and Chapter 7 describes the selections made using these objects. The trigger modeling is discussed in Chapter 8 and the background estimation strategy is described in Chapter 9. Chapter 10 defines the model-independent discovery regions and model-dependent exclusion regions. Chapter 11 describes the systematic uncertainties and how they were estimated. Chapter 12 discusses potential issues arising from signal contamination in the control and validation regions. Chapter 13 presents the results for both model-independent and model-dependent scenarios. Chapter 14 suggests future ideas for the analysis. Part IV presents a search for di-Higgs production targeting modified Higgs self-couplings. Chapter 15 describes my work simulating different models by reweighting the parameter $\kappa_\lambda = \lambda_{hhh}/\lambda_{hhh}^{\mathrm{SM}}$. Chapter 16 discusses the $HH \to b\bar{b}b\bar{b}$ analysis channel while Chapter 17 describes the combination of channels with various final states. Ideas for future di-Higgs analyses are discussed in Chapter 18. Part V gives some concluding remarks.

Chapter 2

THEORY

2.1 The Standard Model

The current, state-of-the-art theory of particle physics is called the Standard Model, or SM [6]. The SM uses a quantum field theory (QFT) framework to describe all known elementary particles, as well as their interactions through 3 of the 4 fundamental forces (electromagnetism, the weak force, and the strong force; but not gravity). It has withstood numerous tests, and describes nearly all known phenomena with astonishing accuracy.

Particles in the Standard Model are excitations of quantum fields. The SM contains 13 types of particles (shown in Figure 2.1) and 3 fundamental forces. These are described in detail below.

Particles can be classified as either fermions or bosons. A fermion is a particle for which the wave function is anti-symmetric under particle exchanges, while a boson is a particle for which the wave function is symmetric under particle exchanges. In particular, this leads to an important property of fermions, called the Pauli exclusion principle: no two identical fermions can occupy the same quantum state at the same time. Conversely, bosons obey Bose-Einstein statistics, and can be found in the same state. It can be shown that bosons have integer spins while fermions have half-integer spins [8].

2.1.1 Fermions

Fermions are the building blocks of matter. There are 3 generations of elementary fermions. Particles in the first generation (the first column in Figure 2.1) are stable or can form stable bound states, while higher generations differ only in the masses of the particles and can decay to lighter generations (plus neutrinos for lepton decays). Each generation can be subdivided into 2 quarks and 2 leptons. In addition to the particles listed below,

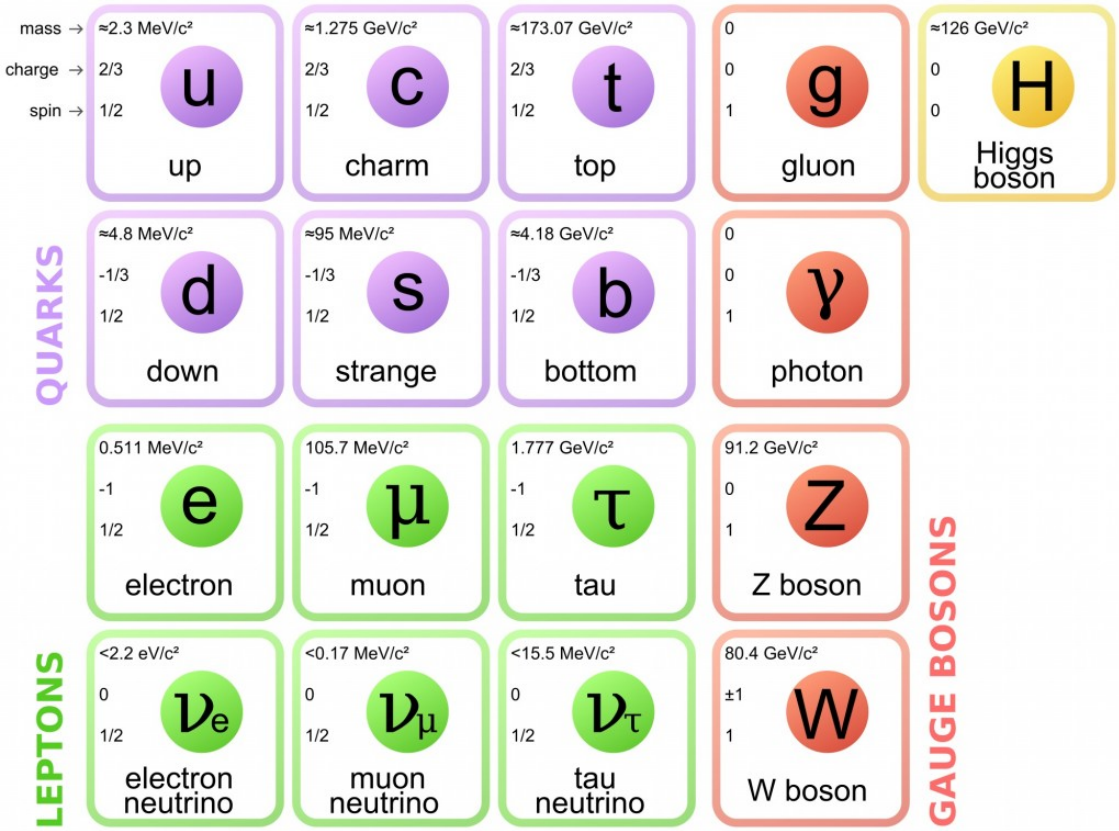


Figure 2.1: A diagram showing the elementary particles of the Standard Model [7].

each particle has a corresponding antiparticle with the same mass but opposite quantum numbers.

One quark (up-type) has electric charge of $+\frac{2}{3}$ while the other (down-type) has charge of $-\frac{1}{3}$ (in units of the electron charge). All quarks carry one unit of color charge, which can be either red, green, or blue (or anti-red, anti-green, or anti-blue). Quarks are also charged under the weak interaction. The quarks are named "up" and "down" for the first generation, "charm" and "strange" for the second generation, and "top" (or "truth") and "bottom" (or "beauty") for the third generation.

The leptons can be subdivided into charged leptons and neutrinos (neutral leptons). The charged leptons have electric charge of -1, neutral color charge, and are charged under the weak interaction. These are the electron, muon, and tau. Each has a corresponding type of neutrino, which is charged only under the weak interaction. Neutrinos only exist in left-chiral states, while anti-neutrinos only exist in right-chiral states.

The neutrinos are notable for their masses, which are orders of magnitude lower than that of their corresponding charged leptons. In addition, the weak eigenstates (electron, muon, and tau neutrino) do not correspond to the mass eigenstates. This leads to a phenomenon called neutrino oscillation, where a freely propagating neutrino will oscillate between the three flavors [9]. The mixing angles are given by the unitary Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [10]

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} 0.801 - 0.845, & 0.513 - 0.579, & 0.143 - 0.155 \\ 0.234 - 0.500, & 0.471 - 0.689, & 0.637 - 0.776 \\ 0.271 - 0.525, & 0.477 - 0.694, & 0.613 - 0.756 \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{2.1}$$

where each entry shows the range of values consistent with experiment and ν_1, ν_2, and ν_3 are the mass eigenstates.

Force	Strength
Strong	> 1
Electromagnetic (α)	1/137
Weak (α_w/M_W^2)	$5*10^{-12}$
Gravity	10^{-42}

Table 2.1: The relative strengths of the fundamental forces at low energies. Note that the values are dependent on the energy scale. The weak force has a coupling constant of $\frac{1}{29.5}$ but is suppressed by the W mass [3].

2.1.2 Gauge Bosons

Bosons (with the exception of the Higgs) carry the fundamental forces, with relative strengths shown in Table 2.1. There are four such types of gauge boson, each with a spin of 1. Each boson has an antiparticle with the same mass but opposite quantum numbers, except for the photon and Z boson, which are their own antiparticles.

The theory of the strong force is called quantum chromodynamics (QCD). QCD interactions are mediated by gluons. These are 8 massless, neutral particles. As each carries a unit of color charge and a unit of anti-color charge, gluons cans interact with each other. The strong force is responsible for binding quarks together to form protons and neutrons. Particles with color charge (quarks and gluons) are collectively referred to as partons [3].

As its name suggests, the strong force is the strongest of the fundamental forces. Furthermore, its strength increases at large distances. Due to this, quarks exhibit a property known as "confinement". This means that they can only occur in colorless (the strong force equivalent of neutral) combinations. This occurs because the energy required to pull a quark out of a bound system is sufficient to create new quarks. These new quarks are created in such a way so as to create colorless bound states, with the reduction in potential energy of the previously free quarks being enough to make the process energetically favorable. These bound states are called hadrons and can be either mesons, containing two quarks of opposite

color (such as $r\bar{r}$), or baryons, containing three quarks of complementary color (such as rgb). Conversely, the strong force weakens at short distances (high energy). This property is called asymptotic freedom [3]. As gluons are massless, the strong force theoretically has infinite range. However, confinement leads to a limited effective range.

The electroweak bosons arise from two components: a triplet of bosonic fields W_μ^i that couple to the weak isospin $\mathbf{I}$ and a singlet B_μ that couples to the weak hypercharge Y. These are related by the formula

$$Q = I_3 + \frac{1}{2}Y \tag{2.2}$$

where Q is the electric charge. The first two components of W_μ mix to form charged bosons called the $W^\pm$, given by

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}} \tag{2.3}$$

The remaining component mixes with the singlet B_μ as

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \sin\theta_w & \cos\theta_w \\ \cos\theta_w & -\sin\theta_w \end{pmatrix} \begin{pmatrix} W^3 \\ B \end{pmatrix} \tag{2.4}$$

to produce the massless photon γ and the massive neutral weak boson Z. θ_w is the weak mixing angle and has a measured value of 28.2° [11]. The photon carries the electromagnetic force, while the $W^\pm$ and Z bosons carry the weak force [3]. The electromagnetic (EM) force describes the interactions between electrically charged particles. This is responsible for most mundane encounters: when you grab an object, the repulsion between the electrons in the object and those in your hand prevent your fingers from passing through it. As the photon is massless, the electromagnetic force has infinite range. This is governed by the classic $\frac{1}{r}$ Coulomb potential.

The weak force governs interactions between fermions based on their weak hypercharge. The fundamental strength of the force is larger than that of electromagnetism but is suppressed by the large masses of the $W^\pm$ and Z boson mediators. The $W^\pm$ bosons govern the

charged current interaction while the Z boson governs the neutral current interaction. The charged current interaction only affects left-chiral particles (and right-chiral antiparticles).

The quark eigenstates of the weak nuclear force are mixed states with respect to the strong eigenstates used to define the particles. The transformation is given by the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix [11].

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} 0.97370 & 0.2245 & 0.00382 \\ 0.221 & 0.987 & 0.0410 \\ 0.008 & 0.00388 & 1.013 \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.5}$$

where d', s', and b' are the weak eigenstates while d, s, and b are the strong (and mass) eigenstates. The up-type quarks do not need to be rotated as the strong and weak eigenstates match.

This allows the weak interaction to violate flavor symmetries. While the electromagnetic and strong forces always preserve the number of particles in each generation (with anti-particles counting as negative), the charged current interaction allows inter-generational decays. This prevents the second and third generation quarks from forming stable hadrons.

The remaining force, gravity, is not described by the Standard Model.

2.1.3 Conservation Laws

The Standard Model has several conserved quantities. The continuous symmetries are energy, momentum, angular momentum and Lorentz invariance. A beautiful result known as Noether's Theorem shows that these each correspond to a conserved quantity. Energy is conserved because the laws of physics are the same at any time while momentum is conserved because the laws of physics are the same at any point in space. The invariance of physics under rotations leads to conservation of angular momentum.

There are also conserved discrete symmetries. Baryon number (B) is $1/3$ for quarks and $-1/3$ for anti-quarks. Lepton number (L) is $+1$ for all leptons and -1 for all anti-leptons.

These are not quite conserved due to non-perturbative quantum anomalies, but the difference $B - L$ remains conserved [12].

Three related discrete symmetries play a fundamental role in the Standard Model. Parity symmetry (P) states that the laws of physics are invariant under reflections. Charge conjugation (C) transformations convert particles to their antiparticles. Finally, time reversal symmetry (T) holds when the laws of physics work the same going backward in time. Each of C, P, and T are conserved by the electromagnetic and strong forces but violated by the weak force. However, the combination of C and P is nearly conserved by the weak force, and the full CPT symmetry holds in all cases. This last result, called the CPT theorem, leads to the equal masses between particles and their antiparticles [3].

Finally, each type of charge (electric for EM, color for QCD, and weak hypercharge/ isospin for the weak force) is conserved. These correspond to global gauge symmetries, where the SM fields are invariant under global phase transformations. The Standard Model is more demanding, promoting each of these global symmetries to a local symmetry. This allows each interaction to be associated with a group symmetry. The strong force is represented by the special unitary group $SU(3)$, leading to the $3^2 - 1 = 8$ observed gluons of the Standard Model. Due to their mixing, the weak and electromagnetic forces are represented together by $SU_L(2) \times U(1)$, where the L signifies that $SU(2)$ only couples to left-chiral particles. This generates $2^2 - 1$ (from $SU(2)$) $+1$ (from $U(1)$) mediators, corresponding to the W_i and B that combine to form the $W^{\pm}$ bosons, Z boson, and photon [3].

2.1.4 Higgs Mechanism

Under the above formalism, gauge invariance requires that the fermions and gauge bosons be massless. This is in contradiction to observations, where the $W^{\pm}$ bosons, Z bosons, and fermions are all massive. This problem is resolved by the introduction of the Higgs boson.

Proposed by Peter Higgs, François Englert, Robert Brout, Gerald Guralnik, C. Richard Hagen, and Tom Kibble in [13–16], the Higgs boson is a complex scalar doublet $\Phi =$

Conserved Quantity	Symmetry
Energy	Translations in time
Momentum	Translations in space
Angular momentum	Rotations
The product CPT	Charge-parity and time reversal
Charge (electric, color, and weak separately)	Global gauge transformations

Table 2.2: A selection of conserved quantities in the Standard Model and their corresponding symmetries [3].

$(\Phi^+, \Phi^0)^{\mathrm{T}}$ with potential

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{2.6}$$

where $\mu^2 = m_{\mathrm{H}}^2/2$ is the mass term and λ is a constant. This is shown in Figure 2.2. The minimum of this potential does not occur at Φ=$\mathbf{0}$, but rather at $\Phi = \frac{1}{\sqrt{2}}(0, \sqrt{\mu^2/2\lambda})^{\mathrm{T}}$. The value $\sqrt{\mu^2/2\lambda}$ is known as the vacuum expectation value v and has been measured to be 246 GeV [11]. We observe the originally symmetric field about the non-zero vacuum expectation value, breaking the symmetry. By transforming to coordinates centered around the vacuum expectation value, we generate 4 scalar bosons (one for each degree of freedom in the Higgs doublet). Three of these are massless Nambu-Goldstone bosons while the last is the massive Higgs boson. This process is called spontaneous symmetry breaking [17].

Making a local phase transformation (under which the SM is invariant) allows the Goldstone bosons to be reinterpreted. For the right transformation, these disappear and are replaced by mass terms for the $W^\pm$ and Z bosons from Equations 2.3 and 2.4. The masses are proportional to the couplings between the gauge bosons and the Higgs boson. Furthermore, it can be shown that the resulting Yukawa interactions with the fermions generates masses proportional to their coupling strengths [17].

After decades of effort, the Higgs boson was discovered by the ATLAS and CMS collaborations in 2012 [4, 5]. The observed mass of around 125 GeV can be seen clearly in Figure

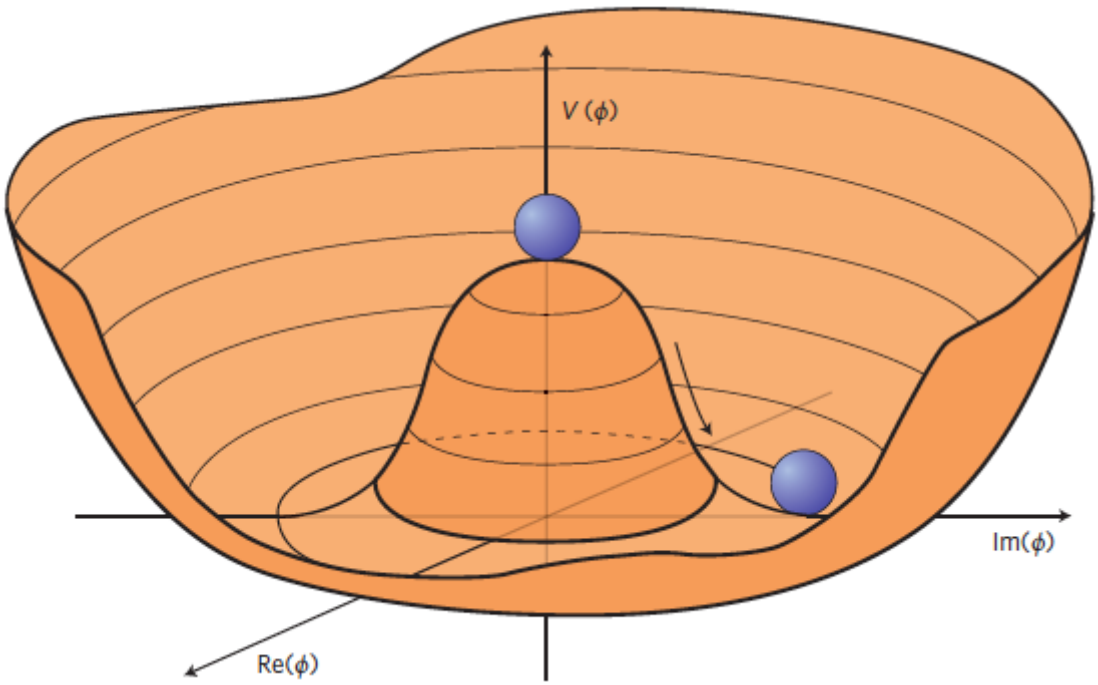


Figure 2.2: A diagram of the Higgs potential [17].

2.3. This has since been updated to 125.10±0.14 GeV [11].

As shown in Figure 2.4, the Higgs boson has many possible decay modes. For the observed mass, the dominant branching ratio (BR) is $H \to b\bar{b}$, occurring 58% of the time. Other notable channels include $H \to \tau^+\tau^-$ and $H \to \gamma\gamma$. While these have lower statistics than the $b\bar{b}$ channel, they allow for cleaner reconstruction.

2.1.5 Problems with the Standard Model

Despite its many successes, the Standard Model suffers from several problems. It cannot account for the existence of dark matter nor the imbalance between matter and antimatter in the universe. It also fails to explain the observed Higgs boson mass in a natural way. Addressing these issues will require new theories that must go beyond the Standard Model while preserving its accuracy at describing other phenomena. While many such theories have been proposed, we have yet to observe any definitive evidence for them.

The Higgs boson provides a promising tool in these searches. Many new theories predict

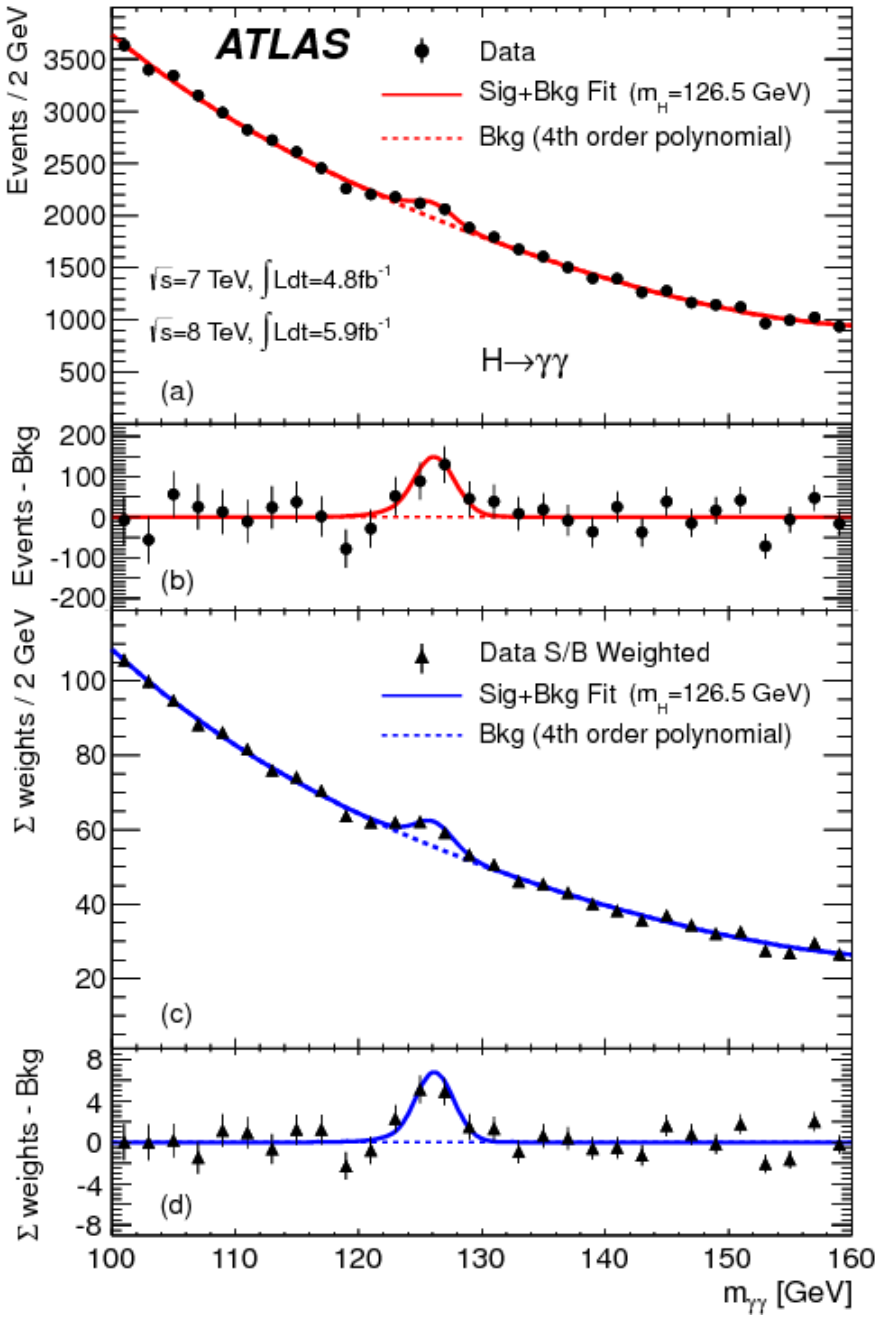


Figure 2.3: Unweighted and weighted fits in the $H \to \gamma\gamma$ channel showing the discovery of the Higgs boson [4].

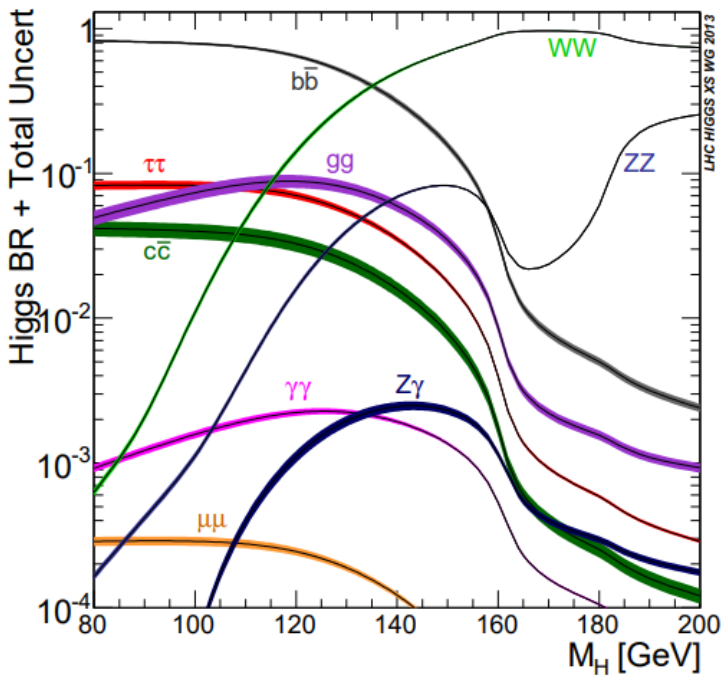


Figure 2.4: Theoretical branching ratios for Higgs boson decays under various mass hypotheses. Note that the Higgs boson only interacts with photons and gluons through loop diagrams [18].

that the production or decay mechanisms for the Higgs boson differ from the Standard Model values. By searching for such deviations, we can either find or constrain new physics.

2.2 Supersymmetry

2.2.1 The Hierarchy Problem

The $\mu^2 = m_h^2/2$ term in Equation 2.6 consists of two components:

$$m_h = m_{\text{bare}} + m_{\text{corr}} \tag{2.7}$$

The first term is the bare mass of the Higgs boson, which is a free parameter in the Standard Model. The second term is the corrections due to loop effects such as the ones shown in

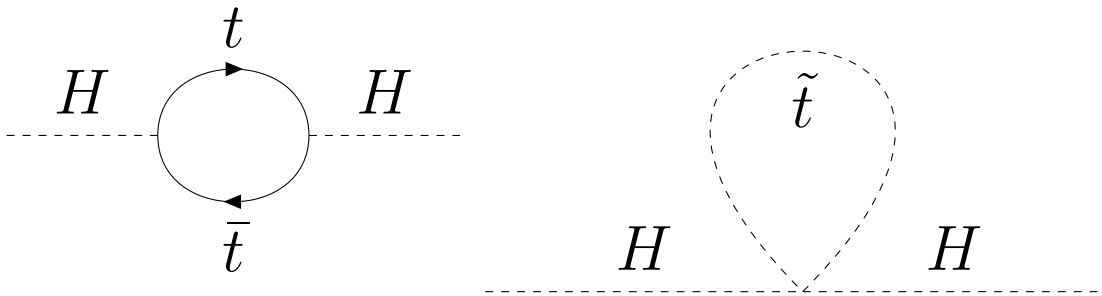
Figure 2.5a. It can be shown that fermions contribute positive corrections of

$$\Delta m_h^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda^2 \tag{2.8}$$

where λ_f is the coupling of the Higgs boson to the fermion in question and Λ is the cut-off energy beyond which the theory is no longer true [19]. Only the highest order term in Λ is shown.

As the heaviest Standard Model particle, the top quark has the largest coupling to the Higgs boson with $\lambda_t \approx 1$. The effects of quantum gravity are expected to break the SM at the Planck scale of $\Lambda \approx 10^{19}$ GeV. If the SM is to hold until this point, then Equation 2.8 implies $m_{\text{corr}} \approx 10^{19}$ GeV. Given the observed Higgs boson mass of 125 GeV, it follows from Equation 2.7 that $m_{\text{bare}} \approx 125 - 10^{19}$ GeV.

This poses a problem for the Standard Model. If m_h is truly a free parameter of order 10^{19} GeV, then it would require a remarkable coincidence for Equation 2.7 to hold. While it is technically possible for this to be the case, the high degree of precision required in the cancellation seems rather unnatural. This conundrum is called the hierarchy problem. The response of many physicists has been to try to develop a theory where the Higgs boson is forced to have a small mass in a natural way. Enter supersymmetry (SUSY) [19].



(a) Higgs self-energy top quark loop. (b) Higgs self-energy stop quark loop.

Figure 2.5: Self-energy corrections to the Higgs boson mass.

SM Particle Name	Superpartner Name
Electron e	Selectron $\tilde{e}$
Neutrino ν	Sneutrino $\tilde{\nu}$
Lepton	Slepton
Fermion	Sfermion
Top quark t	Top squark or stop quark $\tilde{t}$
Gauge boson	Gaugino
Higgs boson h	Higgsino $\tilde{H}$
Graviton G	Gravitino $\tilde{G}$
Particle	Sparticle

Table 2.3: Illustrative examples for SUSY naming conventions.

2.2.2 Supersymmetric Partners

The idea at the heart of SUSY is rather simple: for each known fermion (boson), there is a "superpartner" boson (fermion) with the same quantum numbers, except for spin. This works because scalar boson loops (see Figure 2.5b) contribute negative corrections to m_{corr} of

$$\Delta m_h^2 = \frac{\lambda_S}{16\pi^2}\Lambda^2 \tag{2.9}$$

where λ_S is the coupling term between the Higgs boson and the scalar [19]. As before, only the highest order term in Λ is shown. From Equations 2.8 and 2.9, it follows that if each fermion is partnered with two scalar bosons (or one complex scalar boson) with $\lambda_S = |\lambda_f|^2$, the largest corrections will cancel. Likewise, any contributions from bosons will be nearly canceled by their superpartners.

This describes the simplest possible version of SUSY, the Minimal Supersymmetric Standard Model (MSSM). Each SM fermion has a spin-0 complex scalar boson as a superpartner, with the left- and right-chiral components of the fermions treated separately. The SM spin-1

gauge bosons have spin-1/2 fermion superpartners. In addition, the spin-2 graviton has a spin-3/2 superpartner. The superpartners of fermions are named by appending an "s-" to the beginning of the SM particle name while the superpartners of bosons have "-ino" added to the end. A tilde is placed over particle symbols to represent their supersymmetric counterparts. A list of naming conventions is shown in Table 2.3.

The Higgs boson presents a special case. Supersymmetry requires a minimum of 2 Higgs doublets. The first, H_u, has weak hypercharge $Y = +1/2$ and couples to up-type (up, charm, and top) quarks. The second, H_d, has $Y = -1/2$ and couples to down-type (down, strange, and bottom) quarks as well as leptons [19, 20]. Such a model is called a two Higgs doublet model (2HDM) and creates 8 degrees of freedom in the Higgs states. As usual, three are used to provide the masses of the $W^{\pm}$ and Z bosons. This leaves 5 Higgs bosons made of linear combinations of H_u and H_d: h, H, A, and $H^{\pm}$. The SM Higgs boson is either h or H, with H heavier by convention. A differs in that it is CP-odd, while $H^{\pm}$ are charged. These models are often characterized by

$$\tan\beta = v_u/v_d \tag{2.10}$$

where v_x is the vacuum expectation value for the H_x Higgs boson [19].

The superpartners of the Higgs bosons (higgsinos) mix with the winos and bino. This results in 4 neutral gauginos, called neutralinos ($\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, $\tilde{\chi}_3^0$, $\tilde{\chi}_4^0$) and two charged gauginos (charginos; $\tilde{\chi}_1^{\pm}$, $\tilde{\chi}_2^{\pm}$).

If supersymmetry were a perfect symmetry, all sparticles would have the same properties (except spin) as their particle superpartners. In particular, the masses would be equal. However, experimental evidence has ruled out SUSY particles at such light scales. It follows that if SUSY exists, it must be a broken symmetry. This can be achieved through spontaneous symmetry breaking, similar to the Higgs mechanism described in Subsection 2.1.4. This would generate a Nambu-Goldstone fermion called the goldstino. This is absorbed by the gravitino, providing it with a non-zero mass.

In addition to solving the hierarchy problem, supersymmetry also has the potential to

provide a dark matter candidate. If we define R-parity P_R as

$$P_R = (-1)^{3(B-L)+2s} \tag{2.11}$$

where B is the baryon number, L is the lepton number, and s is the spin, then all Standard Model particles and the additional Higgs bosons will have $P_R = +1$ while their superpartners will all have $P_R = -1$. If R-parity is conserved, then the lightest supersymmetric particle (LSP) will be stable. If the LSP is neutral under the electromagnetic and strong forces, it could provide a WIMP (weakly-interacting massive particle) dark matter candidate. Most R-parity conserving models have either the gravitino $\tilde{G}$ or the lightest neutralino $\tilde{\chi}_1^0$ as the LSP, though the $\tilde{\chi}_1^0$ would be a stronger dark matter candidate [19].

2.2.3 The μ Problem

The vacuum expectation values of H_u and H_d must satisfy

$$v_u^2 + v_d^2 = v^2 \tag{2.12}$$

where v is the SM Higgs vacuum expectation value and v_u and v_d are the vacuum expectation values for H_u and H_d respectively. Given this constraint, one can derive that

$$m_Z^2 \approx -2\mu^2 - 2m_{H_u}^2 \tag{2.13}$$

in the $v_u \gg v_d$ limit or

$$m_Z^2 \approx -2\mu^2 - 2m_{H_d}^2 \tag{2.14}$$

in the $v_u \ll v_d$ limit, where μ is a supersymmetric higgsino mass parameter, m_{H_u} is a supersymmetry-breaking term for the up-type Higgs mass, and m_{H_d} is a supersymmetry-breaking term corresponding for the down-type Higgs mass [19, 21, 22]. This presents a

similar issue as the hierarchy problem: μ and m_{H_u} (or m_{H_d}) must either be of the order of the Z mass or have unnatural cancellations. For a natural SUSY model, we can then predict that μ (and therefore the higgsino masses) are close to the electroweak scale.

2.2.4 Gauge-Mediated Supersymmetry Breaking

There are different models of supersymmetry based on the method by which it is broken. Three prominent examples are Plank-scale-mediated supersymmetry breaking (PMSB), gauge-mediated supersymmetry breaking (GMSB), and anomaly-mediated supersymmetry breaking (AMSB). In PMSB models, SUSY is broken mainly by gravitional interactions, while in GMSB models it is broken by gauge interactions. AMSB models break SUSY through supergravity [19].

The SUSY search in this thesis uses GMSB models [23], which can be placed into the framework of general guage mediation (GGM). These models predict that the gravitino is nearly massless and thus the LSP. In addition, the gravitino absorbs the goldstino, allowing it to be produced via particle decays at timescales relevant to the ATLAS experiment. Many of these models, including the one presented here, have the lightest neutralino $\tilde{\chi}_1^0$ as the next-to-lightest supersymmetric particle (NLSP) [19]. In particular, this search investigates models where the lightest gauginos are higgsino-dominated. These gauginos are thus alternatively referred to as higgsinos and written as $\tilde{H}$. This leads to nearly degenerate masses for the $\tilde{\chi}_1^0$, $\tilde{\chi}_1^\pm$, and $\tilde{\chi}_2^0$ [24].

Within these models, there are three possible decays for the higgsino: $\tilde{H} \to \gamma\tilde{G}$, $\tilde{H} \to Z\tilde{G}$, and $\tilde{H} \to h\tilde{G}$. The branching fraction of the photon decay decreases with higgsino mass. For cases where the higgsino is heavier than the Higgs boson, the other two modes dominate, with the split heavily dependent on the higgsino mass, $\tan\beta$, and other details of the model [25].

The analysis presented here focuses on higgsinos of with relatively low masses ($\lessapprox$ 300 GeV) but still above the Higgs boson mass. The Feynman diagram for the model under

investigation is shown in Figure 2.6. The gray dot represents the sum of various processes producing the two $\tilde{\chi}_1^0$ sparticles, including through decays of $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$. Other products of the decays are ignored, as the nearly-degenerate mass spectrum of the electroweakinos would make them too soft to detect. Higgsinos are assumed to only decay via $\tilde{H} \to h\tilde{G}$ or $\tilde{H} \to Z\tilde{G}$, with only the Higgs boson decays explicitly considered. The $b\bar{b}$ Higgs boson decay channel is targeted due to its high branching ratio.

This work extends the results from a previous ATLAS search using 2016 data, corresponding to 24.3 fb^{-1} for the low-mass portion of the analysis. The previous search observed an excess of slightly over 2σ (local) for higgsinos with mass $m_{\tilde{H}} \approx 250$ GeV, failing to rule out these scenarios even for a $\tilde{H} \to h\tilde{G}$ BR of 100% [26]. Though outside the scope of this thesis, the search will be complemented by a dedicated analysis targeting high-mass higgsinos.

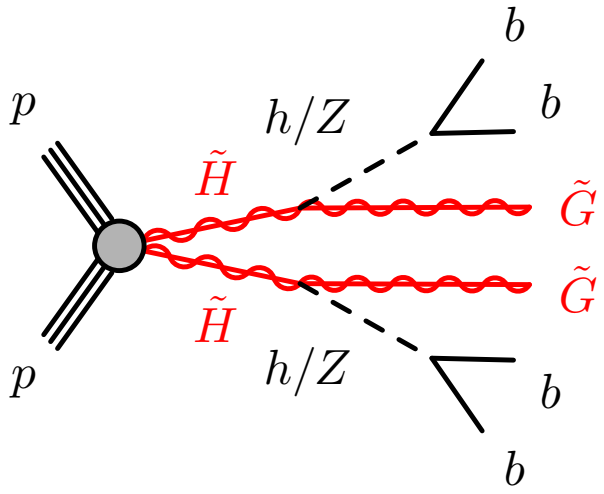


Figure 2.6: Feynman diagram for the GMSB signal model used in this thesis [26].

2.3 Higgs Self-Coupling

By expanding Equation 2.6 about the minimum, the Higgs potential can be written as

$$V(h) = V(0) + \frac{m_h^2}{2}h^2 + v\lambda h^3 + \frac{\lambda}{4}h^4 = V(0) + \frac{m_h^2}{2}h^2 + \frac{m_h^2}{2v}h^3 + \frac{\lambda}{4}h^4 \qquad (2.15)$$

where h is the deviation from the minimum, $v = 246$ GeV is the vacuum expectation value, and $\lambda = \frac{m_h^2}{2v^2} = 0.13$. The h^2 term gives the Higgs boson mass, and has been measured to excellent precision ($m_{\mathrm{h}} = 125.10 \pm 0.14$) [11]. The h^3 term yields the trilinear Higgs self-coupling, with the coefficient giving the coupling strength. Likewise, the h^4 term and coefficient yield the quartic Higgs self-coupling. Notably, the mass scale and vacuum expectation value completely define these coefficients and thus the shape of the potential. Any deviations from the predicted values would be signs of new physics [27].

2.3.1 Motivation

An outstanding problem in the Standard Model is that it does not explain our universe's imbalance of matter and antimatter. The Standard Model predicts that particles and antiparticles should have formed in roughly equal numbers; however, the universe we observe is heavily dominated by matter. The question then arises of how this imbalance came to be. It can be shown that three conditions, called the Sakharov conditions [28], are necessary to achieve this imbalance:

1. Baryon number violation
2. C and CP violation
3. Departure from thermal equilibrium

It is intuitive that there must be processes violating baryon number in order for the baryon number of the universe to grow imbalanced. As discussed in Section 2.1, this only occurs in the Standard Model through non-perturbative quantum anomalies. These are, however, insufficient to explain the observed baryon asymmetry. C and CP violation are required to prevent the conjugate of baryon number violating processes from occurring in equal amounts. While these are violated by the weak nuclear force, the magnitude of the effects are too small to account for our observations. Finally, without a departure from thermal equilibrium, any

process creating additional matter would be canceled by the reverse process, preventing a surplus of matter from accumulating [3].

One possible way of explaining this asymmetry is through electroweak baryogenesis [29, 30]. This hypothesis posits that the imbalance arose during the universe's electroweak phase transition, where the electromagnetic and weak forces separated. However, this would require the transition to be strongly first-order, whereas in the Standard Model the transition is smooth and continuous.

As it arises from electroweak symmetry breaking, the shape of the Higgs potential is sensitive to the details of this phase transition. In particular, changes to the trilinear Higgs self-coupling λ by a factor of 2 can make models of electroweak baryogenesis viable [27].

The shape of the Higgs potential also has implications for the stability of our universe. If the Higgs boson corresponds to the global minimum of the Higgs potential, then the vacuum is stable— it should hold for all time. The other possibility is more troublesome— that the Higgs boson corresponds to only a local minimum, with the true minimum falling at higher energies. If this is the case, the vacuum is said to be meta-stable. In this scenario, the universe could eventually tunnel to the true minimum, changing the laws of physics as we know them.

Current measurements of the top quark and Higgs boson masses imply that the universe is indeed meta-stable (see Figure 2.7) [31]. However, this assumes the Standard Model form of the Higgs potential. Different shapes of the potential, in particular the self-coupling, could place the universe safely in the stable regime [27].

2.3.2 κ Framework

The trilinear term can be investigated through Higgs boson pair production. This is an extremely rare process at the Large Hadron Collider (LHC). The cross sections of the leading production modes are shown in Figure 2.8. At the LHC energy of $\sqrt{s} = 13$ TeV, gluon-gluon fusion (called ggF) dominates with a Standard Model cross section of $\sigma = 31$

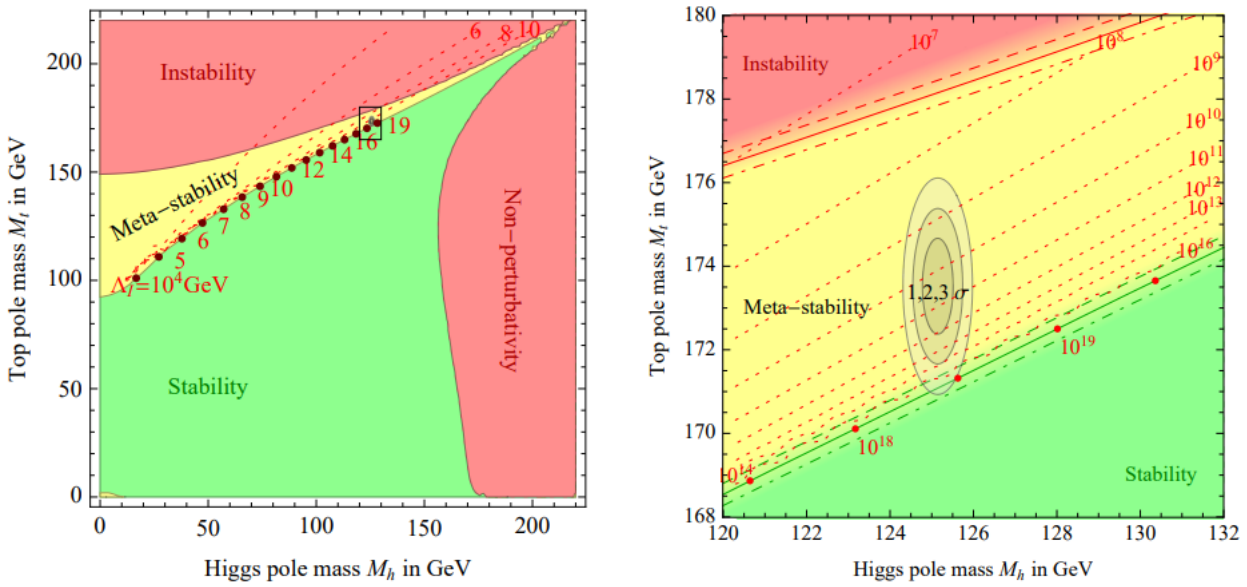


Figure 2.7: Stability of the Standard Model vacuum. The right plot shows a zoomed in view [31].

fb^{-1} [27], with the sub-leading production from vector boson fusion (VBF). At the highest order, two types of processes contribute to ggF production (Figure 2.9). The box diagram produces two Higgs bosons through a loop of four top or bottom quarks, while the triangle diagram produces them through a loop of three top or bottom quarks. These production mechanisms interfere destructively, leading to the low predicted cross section.

The trilinear term can be modeled using the variable κ_λ, defined as $\kappa_\lambda \equiv \lambda_{hhh}/\lambda_{hhh}^{\mathrm{SM}}$. In the SMEFT (Standard Model Effective Field Theory) approach, which treats the Higgs field as a component of an SU(2)xU(1) doublet, this is written as $\kappa_\lambda \equiv \lambda_{hhh}/\lambda_{hhh}^{\mathrm{SM}} \equiv 1 - \frac{3}{2}c_H + c_6$ [27]. As κ_λ varies, the contribution from the triangle diagram changes, strengthening or lessening the effect of the interference. From the LHC-HH working group [32], the ggF HH production cross section is

$$\sigma_{\mathrm{ggF}} = 70.3874 - 50.4111\kappa_\lambda + 11.0595\kappa_\lambda^2 \text{ fb} \tag{2.16}$$

This is shown in Figure 2.10. The cross section is minimized for $\kappa_\lambda \approx 2.3$, but can grow orders of magnitude above the Standard Model value for large $|\kappa_\lambda|$.

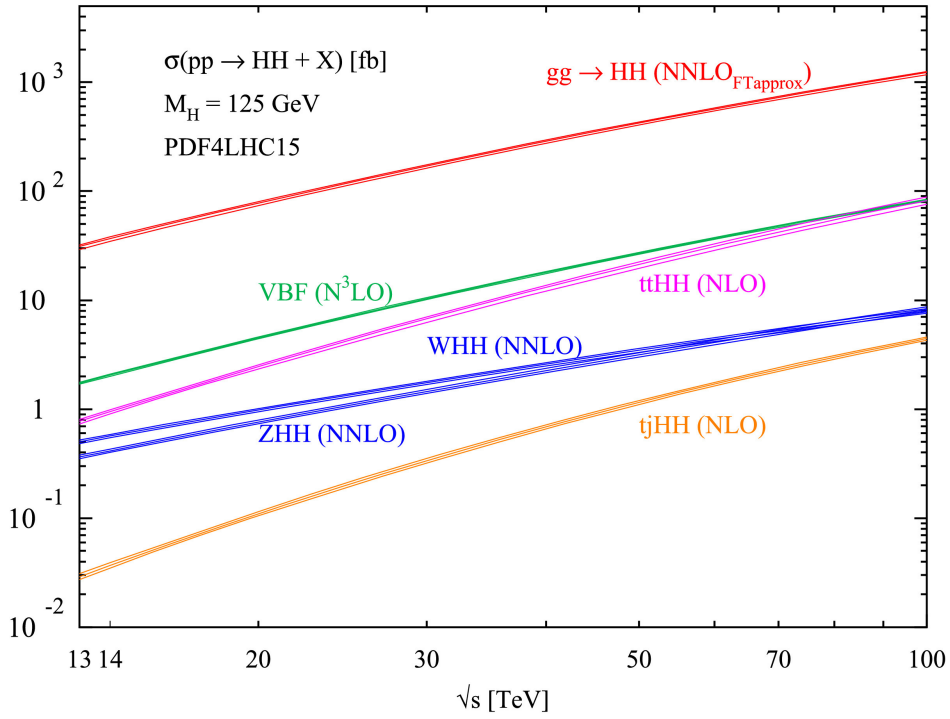
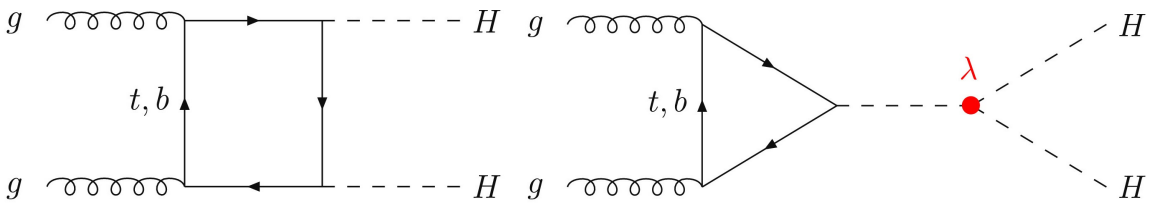


Figure 2.8: Theoretical production cross section for Higgs boson pair production in proton-proton collisions. These assume the Standard Model [27].



(a) Box diagram (b) Triangle diagram.

Figure 2.9: Feynman diagrams for Higgs boson pair production via gluon-gluon fusion [27].

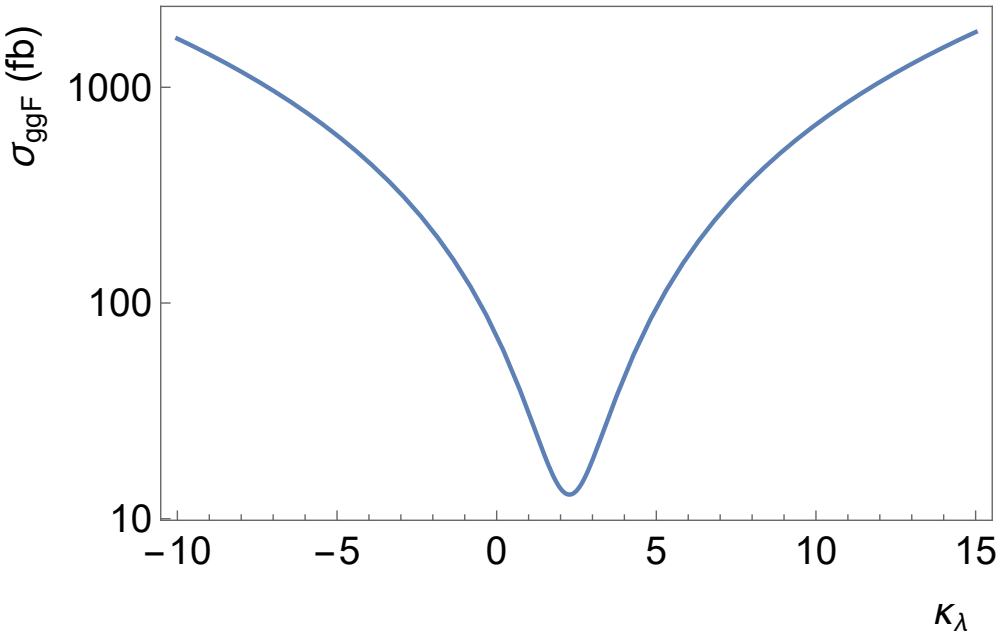


Figure 2.10: Cross section vs. κ_λ for the ggF HH production mode. Numbers are from [32].

Di-Higgs pair production can be investigated in several channels using the different decays of the Higgs bosons. An overview of the channels is shown in Figure 2.11. Of particular note are the $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$, and $b\bar{b}\tau\tau$ channels. The $b\bar{b}b\bar{b}$ channel has the highest branching ratio (33.9%), but suffers from a large QCD multijet background. The $b\bar{b}\gamma\gamma$ channel, on the other hand, has a minuscule branching ratio (0.3%) but an experimentally clean signature. Finally, the $b\bar{b}\tau\tau$ channel holds the middle ground, with a moderate branching ratio (7.3%) and moderate backgrounds. This work shows results from the $b\bar{b}b\bar{b}$ and combined channels, extending the results from previous searches [33, 34] with 24.3 fb^{-1} to 36.1 fb^{-1}.

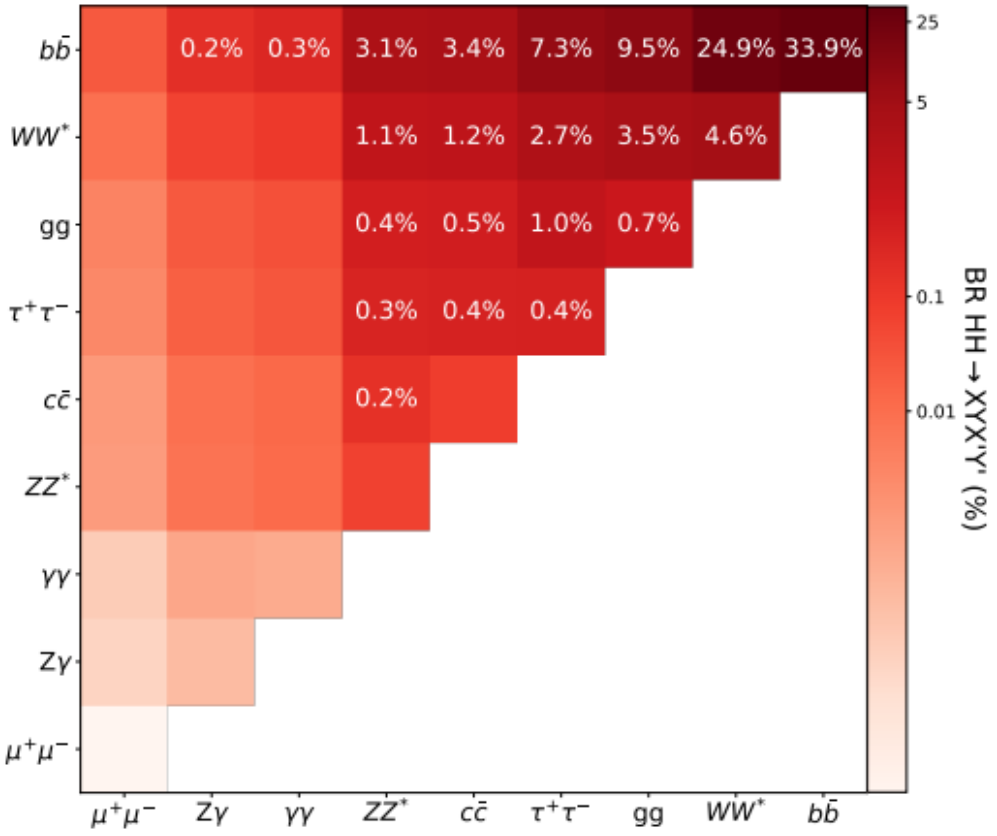


Figure 2.11: Decay channels of Higgs boson pairs and their branching ratios [27].

Part II

The Large Hadron Collider and the ATLAS Detector

Chapter 3

THE LARGE HADRON COLLIDER

3.1 The Large Hadron Collider

In order to search for di-Higgs final states, it is necessary to have a means to create them. The relatively large mass of the Higgs boson (as well as possible SUSY particles) requires a high-energy source. Given the low predicted cross sections for di-Higgs production processes, it is also necessary for the source to produce a large number of events. Both of this criteria are met by the Large Hadron Collider (LHC).

The LHC is a circular particle collider 27 km in circumference [35], making it the largest particle accelerator to date. An aerial diagram is shown in Figure 3.1. Located underground near the French-Swiss border, the LHC accelerates protons to nearly the speed of light. These protons then collide with a center-of-mass energy of 13 TeV. A fraction of this energy is converted to mass through the creation of new particles. Several dedicated experiments study these particles to make precision measurements of the Standard Model as well as to search for new physics.

The LHC collisions occur at four interaction points along the accelerator. Each of these is used by one of the LHC's major experiments. The A Toroidal LHC ApparatuS (ATLAS) and Compact Muon Solenoid (CMS) experiments are general purpose particle detectors designed to search for and study a broad range of physics. The Large Hadron Collider beauty (LHCb) experiment is designed to study the bottom (also called "beauty") quark to search for CP violation . ALICE (A Large Ion Collider Experiment) uses collisions of lead ions instead of protons to study the quark-gluon plasma, an exotic form of matter that was abundant shortly after the Big Bang. Four smaller experiments—TOTEM, LHCf, MoEDAL, and FASER—use detectors near the main experiments for various physics searches [36].

Protons are accelerated using a multi-stage process shown in Figure 3.2. The linear accelerator 4 (Linac4) accelerates hydrogen ions to 160 MeV. The electrons are then removed,

with the resulting protons injected into the Proton Synchrotron Booster (PSB). This accelerates them to 2 GeV [37]. Next, the Proton Synchrotron (PS) brings the protons to 26 GeV and the Super Proton Synchrotron (SPS) accelerates them to 450 GeV. Finally, the LHC accelerates the protons from 450 GeV to their final energy of 6.5 TeV [38].

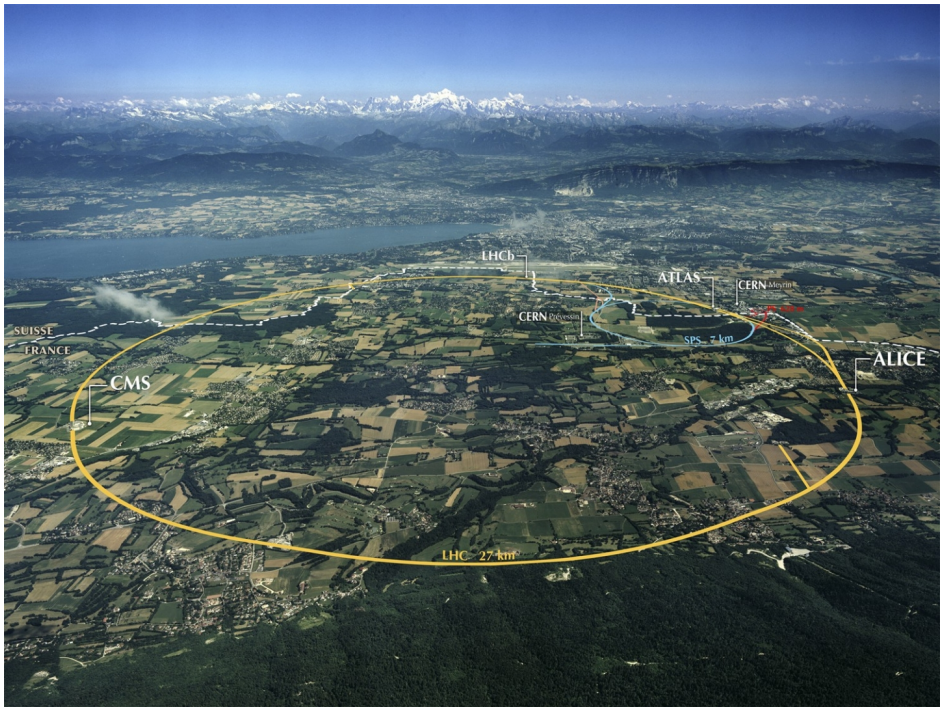


Figure 3.1: An aerial diagram of the LHC. Figure from [39], image credit Maximilien Brice/ CERN.

One of the key properties of the LHC is its high luminosity. The "instantaneous luminosity" is defined as the number of collisions produced per square centimeter per second. Integrating this over time yields the "integrated luminosity" L_{int}. The number of events N produced for a given process is then simply given by the product of the luminosity and cross section

$$N = L_{\text{int}}\sigma_{\text{process}} \tag{3.1}$$

This makes it clear that we want a collider with as high a luminosity as possible. Two main parameters can be adjusted to achieve this. First, protons can be grouped into bunches. The luminosity will then increase with the square of the number of protons per bunch.

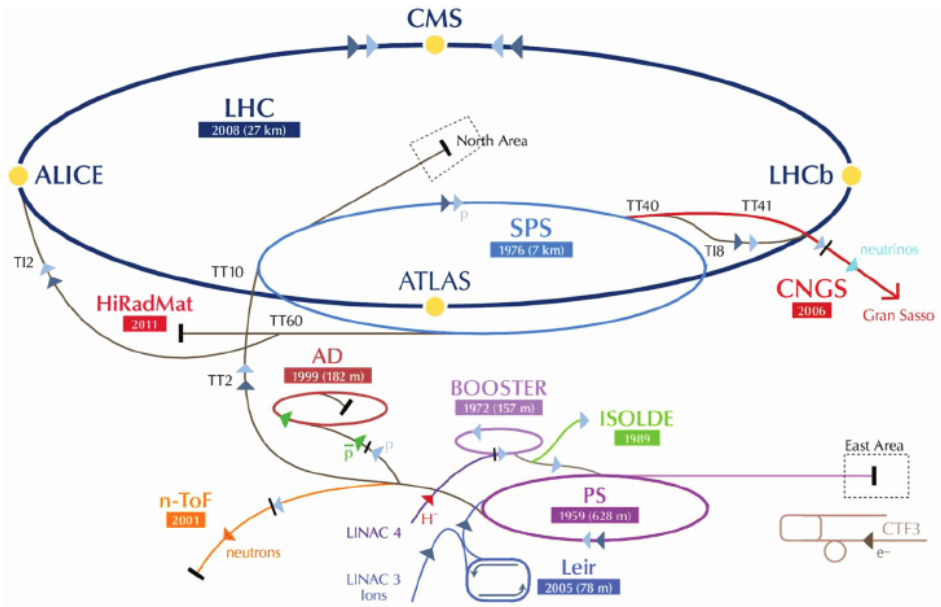


Figure 3.2: The LHC injection chain [40].

Parameter	Value
Circumference	27 km
Collision energy	13 TeV
Collisions/second	1 billion
Peak luminosity	$2.1 * 10^{34}$ cm^{-2}s^{-1}
Protons/bunch	$1.2 * 10^{11}$
Bunch spacing	25 ns

Table 3.1: Summary of LHC properties. Information from [41, 42].

Second, the time between bunches can be shortened, thus increasing the rate of bunch collisions. The LHC has bunches spaced 25 ns apart, each containing 120 billion protons [41]. This has lead to instantaneous luminosities as high as $2.1 * 10^{34}$ cm^{-2}s^{-1} [42]. A summary of the LHC's properties are shown in Table 3.1.

There are, however, drawbacks to a high luminosity. When the bunch size is increased, more protons will interact in each crossing. This makes data-taking and reconstruction challenging. The average number of interactions per bunch crossing, called pile-up, averaged 36.1 over the 2018 data-taking period [42].

3.2 Acceleration

Accelerating protons to 6.5 TeV is no easy task. A simple and illustrative design used for accelerating particles is that of the cyclotron. A cyclotron, shown in Figure 3.3, consists of two semi-cylindrical metal chambers, called "D"s, placed slightly apart. An alternating current is provided so that the polarity of the Ds oscillates at a given frequency. When a particle of charge q and mass m is between the Ds, it will accelerate toward whichever has the opposite charge (within the Ds, no electrical force is felt). Placing the apparatus in a magnetic field B causes the particles to follow a curved path given by

$$\omega = \frac{qB}{mc} \tag{3.2}$$

where ω is the angular frequency [43]. If ω matches the frequency of the alternating current, the polarity will have switched when the particle exits the first D. It will thus be accelerated again, continuing its circular path to the second D. This will occur every half revolution, increasing the particle energy higher and higher [43].

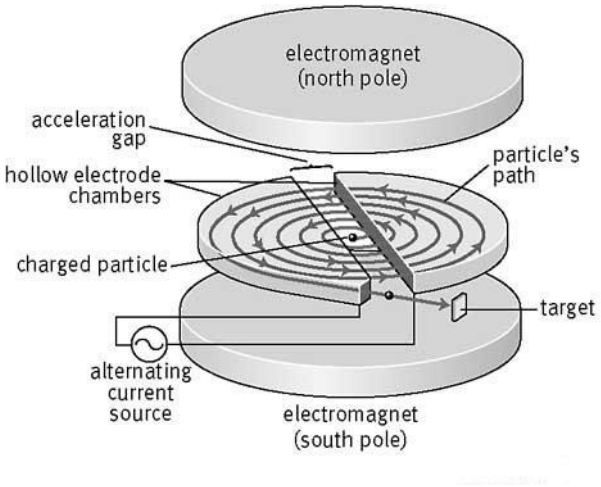


Figure 3.3: Diagram of a cyclotron [44].

A cyclotron cannot, however, reach relativistic energies. This is because Equation 3.2

only holds for non-relativistic particles. The full equation is

$$\omega = \frac{qB}{\gamma mc} \tag{3.3}$$

where γ is the Lorentz factor [43]. The resonance frequency required for continuous acceleration thus varies with the velocity. The LHC overcomes this difficulty by using a more complicated apparatus called a synchrotron.

A synchrotron uses radiofrequency (RF) cavities instead of Ds. These are placed along the beamline and use an electric field to accelerate the particles within them. The LHC uses superconducting RF cavities operating at 400 MHz to provide 485 keV of energy to each proton per revolution [45]. A synchrotron also benefits from self-correcting effects. Consider the case of a collection of particles injected with a small energy spread. From Equation 3.3, lower energy (lower γ) particles will have a higher ω and will thus arrive at the RF cavity early. Likewise, high energy particles will have a lower ω and arrive late. The particles are then out of phase with the RF cavities. If the electric field is decreasing with respect to time at the synchronous phase, a lower energy (earlier arriving) particle will experience more acceleration, bringing it toward the mean energy value. Similarly, higher energy particles will receive less acceleration and be brought toward the mean. While the energy of these particles will continue to oscillate, the energy spread remains small [43, 46].

3.3 Magnets

The LHC makes use of a multitude of magnets to guide the beams and maintain beam stability. The magnets are operated as superconductors, requiring them to be cooled to near absolute zero using liquid helium. These magnets can be divided into several types. The most numerous are the 1232 main dipoles, each 15 meters long [47]. These are used to create an (up to) 8.33 T magnetic field to steer the protons along the circular beamline. A cross-sectional view of a dipole is shown in Figure 3.4. A further 392 quadrupole magnets

are used to focus the beam. For particles that are offset from the beamline on one axis, a properly aligned quadrupole magnet will deflect them toward the center. Particles offset on the other axis will be deflected away from the center. By alternating between quadrupoles focusing in the x-direction and ones focusing in the y-direction, the beam can be constrained to a narrow trajectory (Figure 3.5) [43, 47]. Higher-order magnets (sextupoles, octupoles, decapoles, and dodecapoles) provide additional corrections to the beam path [48].

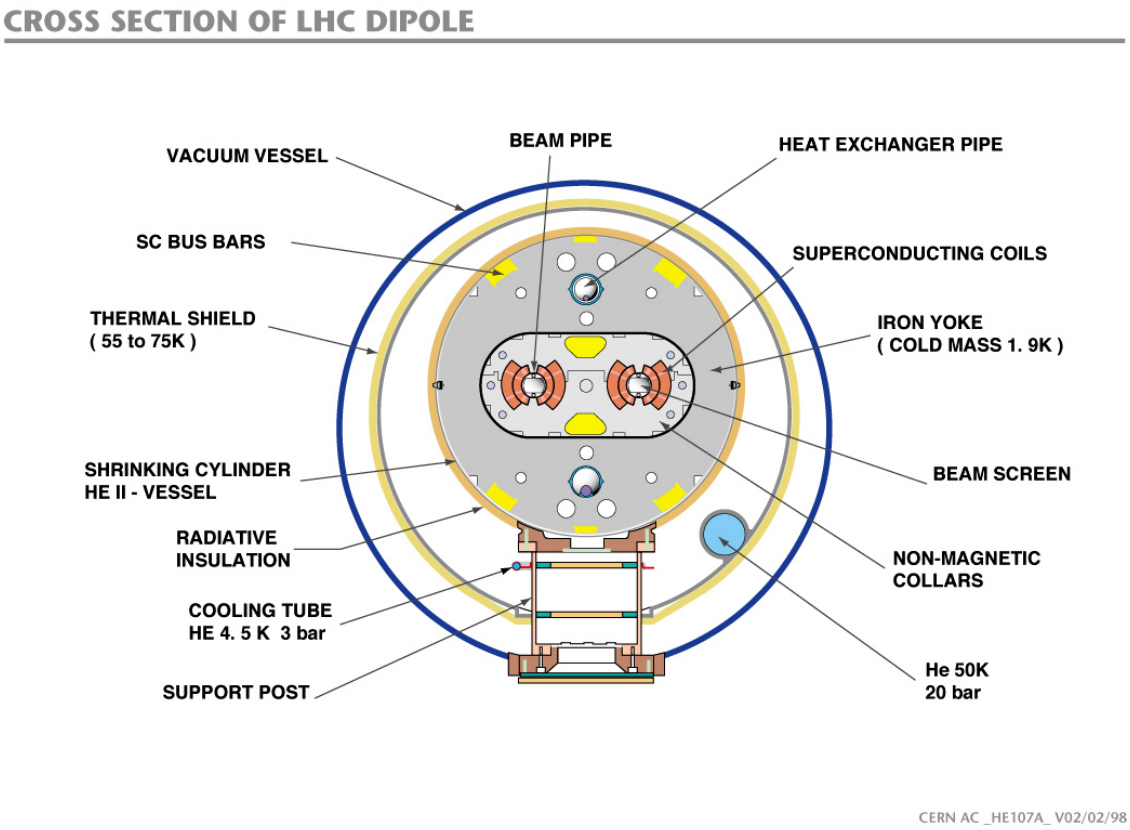


Figure 3.4: A cross section of an LHC dipole [49].

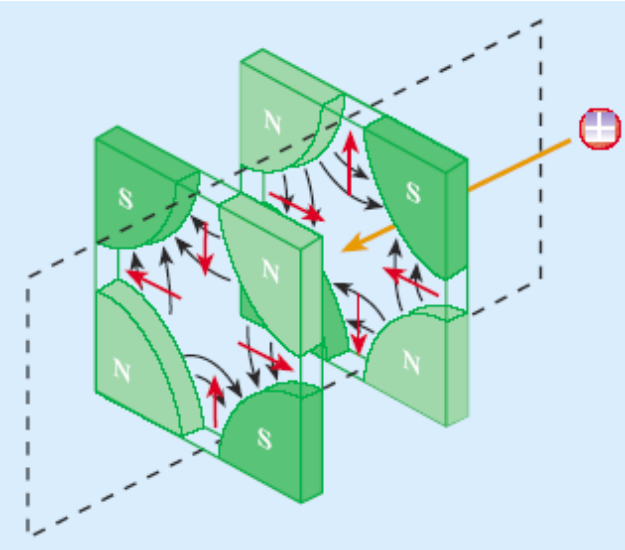


Figure 3.5: A diagram of alternating quadrupole magnets. The deflection of a test particle following the yellow arrow is shown in red [50].

Chapter 4

THE ATLAS DETECTOR

The ATLAS experiment is one of the two general purpose experiments at the Large Hadron Collider. The goal of the experiment is deceptively simple: to detect and measure the particles produced by the LHC in order to study the frontier of high-energy physics. This is primarily done by studying proton-proton collisions at a center-of-mass energy of 13 TeV, though the detector is also used to study heavy-ion collisions at lower energies. This thesis will only cover studies from the proton collision mode.

Achieving this simple goal is a herculean task. The ATLAS detector, shown in Figure 4.1, spans a massive 44m by 25m and weighs 7000 metric tons [51]. The number of personnel working in the collaboration is similarly astounding, featuring over 5000 members from over 40 different countries [52].

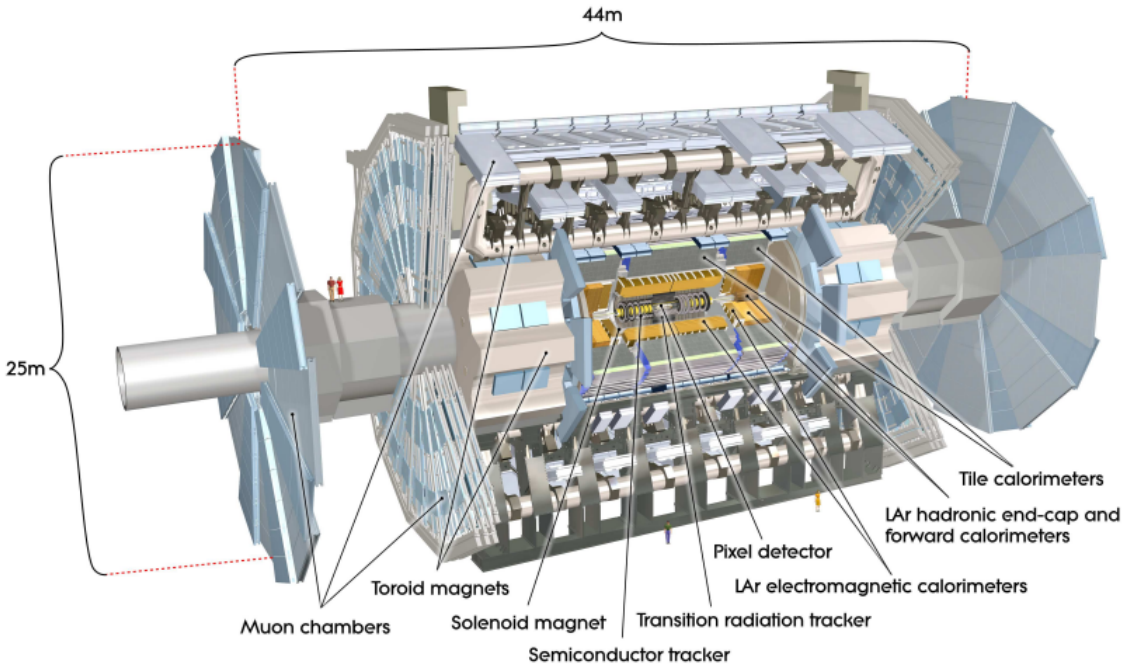


Figure 4.1: A diagram of the ATLAS detector. The size of people are shown for scale on the bottom and left sides of the diagram [51].

The coordinate system used by ATLAS is as follows. The z-axis is defined as the direction of the beamline, with the origin set to the interaction point. The x-axis is perpendicular to

the beamline in the plane of the LHC ring. The direction to the center is defined as the the positive x direction. The y-axis is defined such that up corresponds to +y and down to -y. Alternatively, in cylindrical coordinates, the x-y plane is parameterized using ϕ and η. ϕ is the azimuthal angle while η is the pseudorapidity, given by

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{4.1}$$

where θ is the polar angle. The pseudorapidity is used as differences in η are relativistically invariant for massless particles. In the limit as mass approaches zero, the pseudorapidity approaches the rapidity

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{4.2}$$

The angular distance ΔR between two points is defined as $\sqrt{\Delta\eta^2 + \Delta\phi^2}$ [51].

The ATLAS detector is made up of several sub-detectors, each serving a different purpose. After a collision occurs, outgoing particles are created. The paths of these particles through the detector are illustrated in Figure 4.2. The particles will first hit the Inner Detector (ID). This sub-detector is used to identify, track, and measure the momenta of charged particles. They will then encounter the electromagnetic calorimeters, which provide energy measurements for photons and electrons. Next, the hadronic calorimeters measure the energy of strongly-interacting jets. Finally, the muon spectrometer is used to measure the momenta of muons, which penetrate farther than most other elementary particles. ATLAS is not able to detect neutrinos or certain other neutral particles; their existence must be inferred by balancing the energy of reconstructed events.

4.1 Inner Detector

The Inner Detector is used to track and measure properties of charged particles. More precisely, the goal of the ID is to measure the vertex position, trajectories, and momenta of

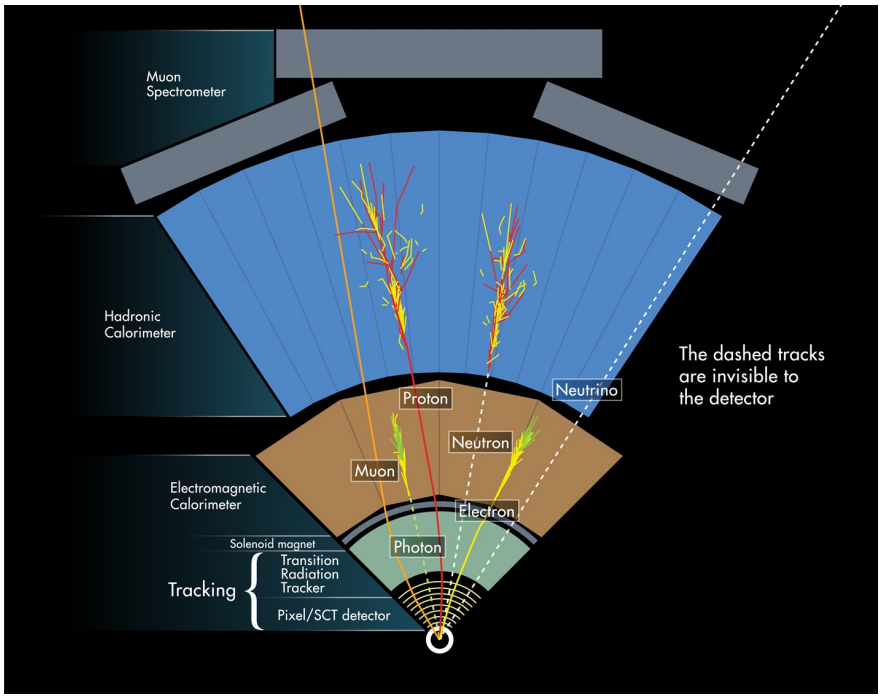


Figure 4.2: A cross-sectional view of the ATLAS detector [53].

charged particles as well as to perform some electron identification. This is accomplished using the cylindrical apparatus shown in Figures 4.3 and 4.4, centered around the interaction point.

The ID consists of 3 parts. Particles produced in a collision will first hit the pixel detectors, then the semiconductor trackers (SCTs), and finally the transition radition trackers (TRT). Each of these has a barrel region corresponding to low $|\eta|$ and an end-cap region corresponding to high $|\eta|$. A superconducting solenoid around the ID generates a 2 T magnetic field.

Together, these sub-detectors provide tracking for charged particles with $|\eta| < 2.5$. The momentum resolution is

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 0.05\%\, p_\mathrm{T} \oplus 1\% \tag{4.3}$$

where $\oplus$ signifies summing in quadrature. In addition, electron identification is available for $|\eta| < 2.0$ [51].

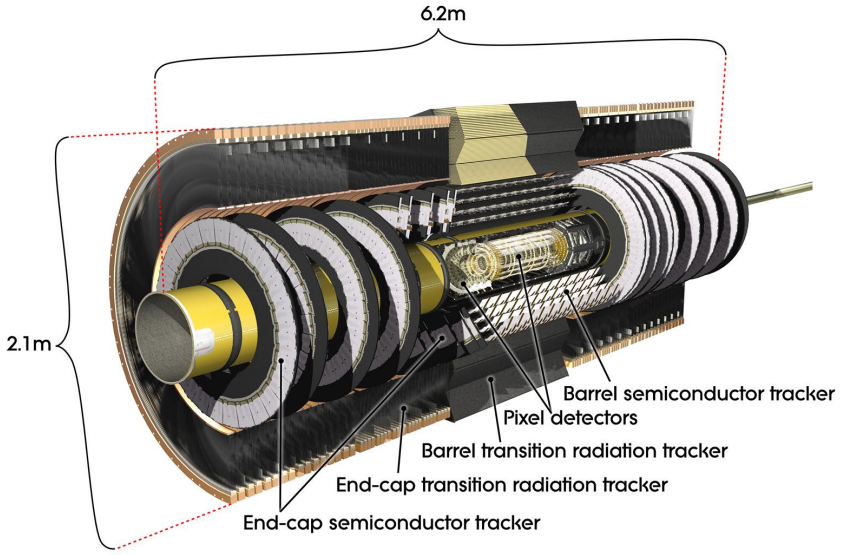


Figure 4.3: A diagram of the ATLAS Inner Detector [51].

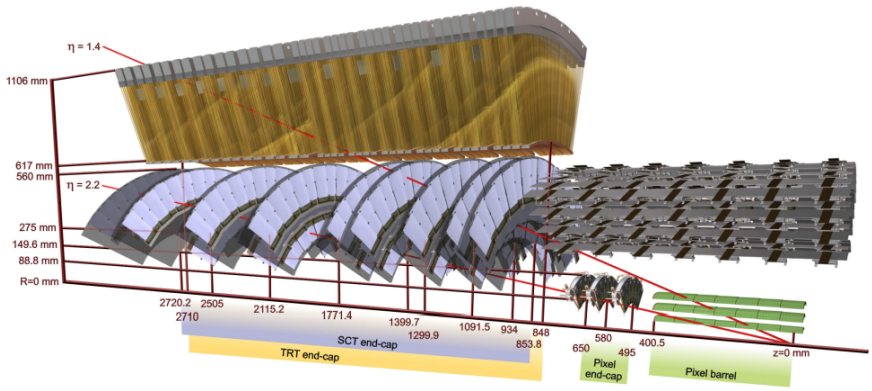


Figure 4.4: A diagram of the ATLAS Inner Detector showing two example particle trajectories [51].

4.1.1 Pixel Detector

The pixel detector consists of 3 cylindrical layers centered on the beamline (in the barrel region), supplemented by a set of 3 disks in each z-direction for the end-cap (see Figure 4.4). Before Run 2, the insertable B-layer (IBL) [54] was installed, providing an extra layer of pixels close to the interaction point. A typical particle will pass through 4 total layers.

Each of the 3 main layers is segmented into pixels with an R–ϕ × z size of 50 × 400 μm^2, providing a resolution of 10 μm in the R–ϕ direction and 115 μm in the z- (barrel) or R- (end-cap) direction. There are a total of 80.4 million pixels in the 3 layers [51]. The IBL layer has a pixel size of 50 × 250 μm^2 R–ϕ × z and provides an additional 12 million pixels [55].

Individual pixel sensors are n-type silicon semiconductor detectors. When excited by an incoming particle, electrons are excited from the valence band of the material to the conduction band. This results in electron-hole pairs. By measuring these pairs, the location of the particle is localized to the sensor [51, 56]. n-type semiconductors have added electron donor impurities, resulting in the production of more conduction electrons per incoming particle at the expense of fewer holes. The total number of charge carriers increases, aiding the detection of the signal [56].

4.1.2 Semiconductor Tracker

The semiconductor tracker, also known as the silicon microstrip (or strip) detector, consists of 4 cylindrical layers in the barrel region plus a set of 9 disks in each z-direction for the end-cap (see Figure 4.4). As the modules constituting these are bi-layers, particles pass through 8 SCT layers [51].

Each SCT bi-layer consists of two 64 mm sensors with readout strips every 80 μm. In the barrel region, one strip is in the z-direction (parallel to the beamline) while the other is at a 40 mrad angle. In the end-cap, one strip is in the R-direction while the other is at a 40

mrad angle. There are a total of about 6.3 million readout strips. The SCT has a resolution of 17 μm in the R–ϕ direction and 580 μm in the z- (barrel) or R- (end-cap) direction [51].

Like the pixel sensors, the SCT sensors use silicon semiconductors. These are setup in the p-in-n configuration [51]. In addition to the bulk n-type region with electron donor impurities, this configuration also has p-type implanted strips electron acceptor impurities [57].

4.1.3 Transition Radiation Tracker

Additional tracking is performed by the Transition Radiation Tracker. The TRT consists of 300,000 drift tubes ("straws"): 50,000 in the barrel region (parallel to the beamline, 144 cm long) and 250,000 in the end-cap (pointing in the radial direction, 37 cm long) [51, 58].

Unlike the pixel and SCT detectors, which extend to $|\eta|$=2.5, the TRT only extends to $|\eta|$=2.0. Also, while the pixel and SCT detectors measure R–ϕ and z to high precision, the TRT only measures R–ϕ, and to a much lower precision of 170 μm. Despite this, the TRT contributes significantly to momentum measurements due to its high number of hits (typically 36 per track) [51, 58].

Each straw of the TRT is a drift tube proportional counter. As a charged particle travels through one of these gas-filled chambers, it ionizes some of the gas molecules. This creates electron-ion pairs. For sufficiently high electric field strengths, a produced electron will have enough energy to interact with a gas molecule and create an additional electron-ion pair. The resulting electrons can do the same; iterating this process creates a Townsend avalanche of many electron-ion pairs. This increases the signal created by a given incoming particle, making it easier to detect. The detector is designed such that the resulting number of pairs is proportional to the number formed in the primary process [56].

The gas used is a 70%/27%/3% mixture of xenon, carbon dioxide, and oxygen. Xenon is needed as the bulk of the mixture due to its x-ray absorption properties, allowing for efficient detection of transition radiation. CO_2 and O_2 acts as quench gases, maintaining the

proportionality of the detector by absorbing photons produced by the ionization processes [58].

4.2 Calorimeters

The ATLAS detector has two types of calorimeters: electromagnetic and hadronic. A diagram of the calorimetry system is shown in Figure 4.5. ATLAS uses sampling calorimeters, which consist of layers of two materials. The first layer is a metallic absorber. Incoming particles traversing this layer are absorbed, creating particle showers in the process. The second layer is either a noble liquid or a scintillator. The particle showers cause noble liquids to ionize and scintillators to emit photons. The resulting signals are then used to reconstruct the energy of the initial incoming particle [59].

The main purpose of the calorimeters is to measure the energy of electrons and photons in the EM calorimeter and hadronic jets in the hadronic calorimeter. This requires that the calorimeters are sufficiently thick to absorb incoming particles of the given type. In addition, they must cover as much of the solid angle around the interaction point as possible to avoid missing any particles. The calorimeters also serve to prevent non-muon events from punching through to the muon spectrometer [51].

The EM calorimeter consists of a presampler, a barrel calorimeter, and two end-caps. The presampler is a thin layer of liquid argon (LAr) spanning $|\eta| < 1.8$. The barrel and end-cap regions both use lead/LAr sandwiches, with the barrel covering $|\eta| < 1.475$ and the end-caps covering $1.375 < |\eta| < 3.2$. The granularity of the calorimeter varies within each region, but values are of the order of 0.025×0.025 $\Delta\eta \times \Delta\phi$ in the barrel region and 0.025×0.1 $\Delta\eta \times \Delta\phi$ in the end-caps [51]. The barrel is > 22 radiation lengths (X_0) thick while the end-caps are > 24 X_0.

The hadronic calorimeter features a barrel (or tile) region covering $|\eta| < 1.7$ and two end-caps covering $1.5 < |\eta| < 3.2$. The barrel uses steel/plastic scintillator sandwiches with

a granularity of (0.1-0.2)×0.1 $\Delta\eta \times \Delta\phi$ [51, 59] and is approximately 7.4 radiation lengths in depth. The end-caps use copper/LAr sandwiches and have a granularity of 0.1×0.1 $\Delta\eta \times \Delta\phi$ up to $|\eta| < 2.5$ and 0.2×0.2 $\Delta\eta \times \Delta\phi$ thereafter.

Additional coverage from $3.1 < |\eta| < 4.9$ is provided by a forward calorimeter in each end-cap. This system consists of one module of copper/LAr sandwiches for EM measurements and two modules of tungsten/LAr sandwiches for hadronic measurements.

An important parameter of the calorimeters is the energy resolution. This determines how precisely the energy of incoming particles can be measured. The resolution can be written as

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{4.4}$$

where $\oplus$ signifies summing in quadrature. The "a" term is caused by stochastic variations in the amount of scintillation light, the "b" term is caused by electrical noise, and the "c" term is due to detector and calibration effects. The EM calorimeter has values of 10%, 170 MeV and 0.7% for a, b, and c respectively, while the hadronic calorimeter has a lower precision of 52.9% and 5.7% for a and c in the barrel region (with negligible contributions from b) [59].

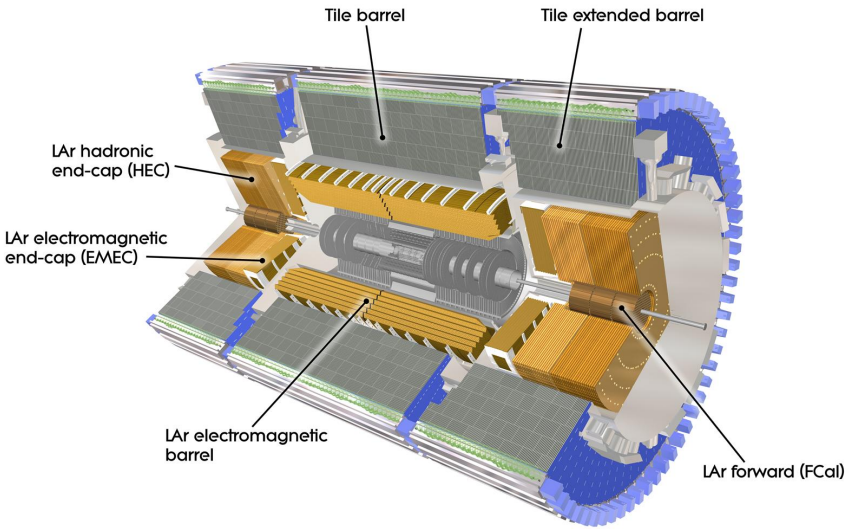


Figure 4.5: A diagram of the ATLAS calorimeters [51].

4.3 Muon Spectrometer

Muons penetrate materials more easily then most elementary particles. It is therefore necessary to have a large, dedicated detector in order to efficiently identify them and measure their momenta. This is accomplished by the muon spectrometer, shown in Figure 4.6.

The muon spectrometer uses 3 large superconducting toroid magnets to deflect muon trajectories. This deflection allows the muons' momenta to be measured. The barrel region has a magnet with a bending power (normal field strength integrated along muon trajectory) of 1.5 to 5.5 Tm while each end-cap has a magnet with a bending power of 1 to 7.5 Tm. Bending is provided by the barrel magnet for $|\eta| < 1.4$, by the end-cap magnets for $|\eta|$ between 1.6 and 2.7, and by a combination of the magnets for $|\eta|$ between those ranges [51].

There are 4 components to the muon spectrometer. The monitored drift tubes (MDTs) and cathode strip chambers (CSCs) are used for tracking while the resistive plate chambers (RPCs) and thin gap chambers (TGCs) are mainly used for triggering. Together, the muon systems provide a transverse momentum resolution of 10% or better for muons with p_{T} up to 1 TeV [60].

The MDT system consists of 2 sets of 3-4 layers of aluminum drift tubes filled with an argon-carbon dioxide gas mixture and cover $|\eta|$ up to 2.0 and 2.7 for the innermost and other layers respectively. The spatial resolution provided by the system is about 50 μm [51, 60].

The cathode strip chambers cover the $2.0 < |\eta| < 2.7$ region for the innermost layer and are comprised of multi-wire proportional chambers (MWPCs) with cathode-strip readout. MWPCs are proportional gas counters with a fine mesh of anode wires. An incoming muon will ionize the gas, causing an avalanche of secondary electrons. These drift toward the anode wires and are picked up, allowing the muon to be localized. The drift tubes comprising the MDTs have similar behavior, but use an additional measurement of the drift time to provide an extra coordinate [11]. The CSCs are used instead of MDTs as they can better handle the high background rate in the region. The CSCs can handle up to about 1000 Hz/cm^2, over

6 times the maximum safe rate for the MDTs [51]. The spatial resolution is about 60 μm in the bending plane [51, 60].

The RPCs cover $|\eta|$ up to 1.05 while the TGCs span $|\eta|$ from 1.05 to 2.7. The RPCs are parallel-plate capacitors placed 2 mm apart. The space between the plates is filled with a gas mixture and subject to a 4.9 kV/mm electric field. The TGCs are multi-wire proportional chambers with a small distance between wires. As with the previous muon spectrometer components, both of these work by measuring the gas ionization from incoming muons [51]. The RPCs and TGCs are used to improve the measurement of ϕ. In addition, they have excellent timing resolution of < 4 ns, allowing them to be used for triggering [60].

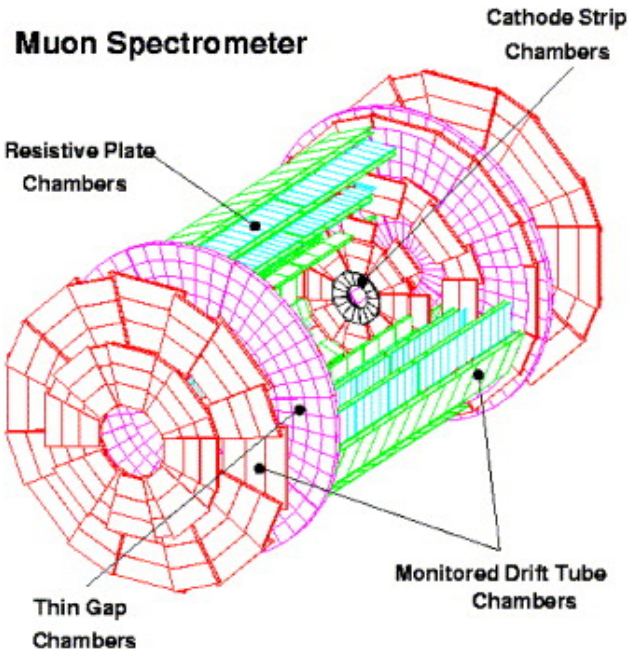


Figure 4.6: A diagram of the ATLAS muon spectrometer [61].

4.4 Jets and Flavor Tagging

Due to quark confinement, when a parton is created by in an LHC collision it will quickly hadronize into a shower of collimated particles. These showers are referred to as jets.

In order to study events, the jets must be reconstructed from their constituent particles.

This is done using the anti-k_t algorithm [62]. This algorithm defines a distance metric between objects i and j of

$$d_{ij} = min(k_{ti}^{-2}, k_{tj}^{-2}) \frac{\Delta\phi^2 + \Delta y^2}{R^2} \tag{4.5}$$

and between object i and the beam B of

$$d_{iB} = k_{ti}^{-2} \tag{4.6}$$

where k_{ti} is the transverse momentum of object i, y is the rapidity of an object, and R is a constant radius parameter. The jet clustering procedure combines objects with their nearest neighbors (in d_{ij}), unless the distance to the beam is smaller than this. This procedure then repeats iteratively using the resulting objects as inputs. This algorithm benefits from insensitivity to soft objects, as the distance to an object will approach infinity as the object's k_t approaches 0.

The anti-k_t algorithm must be seeded with appropriate objects. Two types of objects are used in ATLAS. EMTopo jets, the primary objects for Run 1, used clusters of calorimeter energy deposits. However, this method can be improved. When available, inner detector measurements of object momenta are found to provide more accurate results than the calorimeter energy for low-p_{T} objects. In addition, tracking can be used to reduce the effect of pile-up by identifying whether tracks originated from the primary (hard) vertex. However, neutral particles leave no trace in the inner detector and thus have no tracking information. This dilemma is resolved by the particle flow algorithm. This method begins by matching tracks to calorimeter clusters. The energy expected from tracks is then removed from their associated calorimeter clusters, leaving only the energy from neutral particles. The input objects to the anti-k_t algorithm are then the tracks of charged particles and calorimeter clusters of neutral particles [63].

The dominant decay of the Higgs boson is $h \to b\bar{b}$, occurring 58% of the time. The

analyses in this thesis thus primarily search for Higgs bosons through their decay to b-quarks. This makes the identification of jets arising from b-hadrons (called b-jets) essential. The process of identifying which type of quark a jet arises from is known as flavor tagging.

b-jets have several distinctive features that aid in their identification. First, b-hadrons have a relatively long lifetime on the order of 1.5 ps ($< c\tau > \approx 0.5$ mm) [64]. The hadron thus travels up to a few millimeters before decaying. A diagram comparing prompt jets and a b-jet is shown in Figure 4.7. ATLAS does not directly measure the distance L_{xy} between the primary and secondary vertices. Instead, it uses the transverse impact parameter d_0. This is the distance of closest approach between the primary vertex and the track, if the track were extrapolated backward. b-jets also differ from light- and charm-jets by their relatively high mass, high number of decay products, and fragmentation properties.

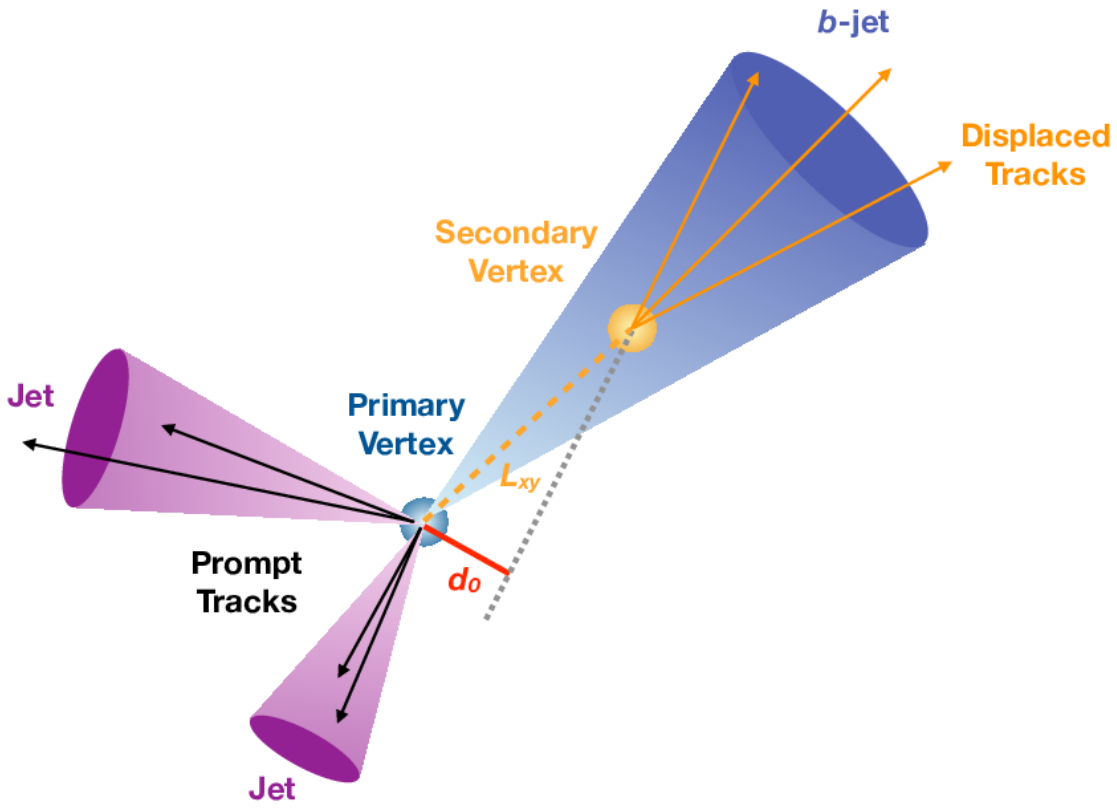


Figure 4.7: A diagram showing the difference between prompt jets and a b-jet. d_0 is the transverse impact parameter [42].

Identifying b-jets is a multi-stage process. First, low-level b-tagging algorithms are run.

These include the RNNIP (Recurrent Neural Network Impact-Parameter) tagger, IP3D, SV1, and JetFitter algorithms. The RNNIP uses a recurrent neural network (RNN) to transform the transverse impact parameter, longitudinal impact parameter, and track/jet information into probabilities of the jet being b, c (charm), τ, or light-flavored. IP3D also uses the impact parameters, but ignores the correlations between tracks in a jet. SV1 reconstructs the secondary vertex and JetFitter reconstructs the b-hadron decay chain [64, 65]. The results of the low-level algorithms are then fed into a deep feed-forward neural network (NN) called DL1r. Combining these with the p_{T} and η of the jets, DL1r assigns probabilities for each jet to be bottom, charm, and light-flavored. These are used to calculate the discriminant

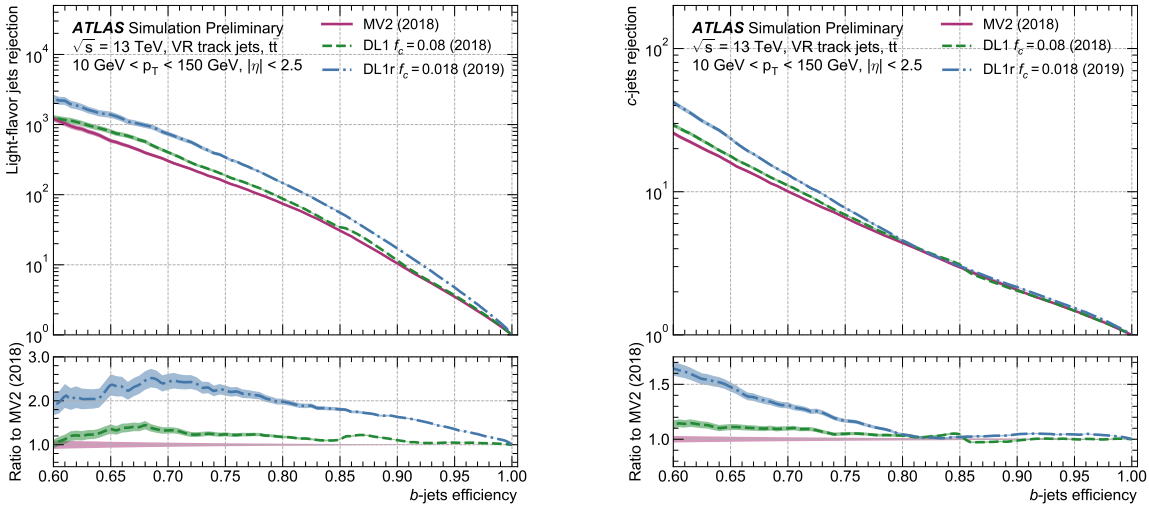
$$D_{\mathrm{DL1r}} = \ln\left(\frac{p_b}{f_c p_c + (1 - f_c) p_{\mathrm{light}}}\right) \tag{4.7}$$

where p_x is the probability for a jet to be x-flavored and f_c is a constant that optimizes the c-jet vs. light-jet rejection. An f_c value of 0.018 was used [64, 66].

Given the score from Equation 4.7, we can set a threshold score above which jets are tagged as b-jets. The optimal value depends on the analysis and is determined by the trade-off between b-jet efficiency and c- and light-jet rejection. Curves showing this trade-off are shown in Figure 4.8. Four working points are defined using the b-jet efficiencies of 60%, 70%, 77%, and 85% respectively.

4.5 Trigger and Data Acquisition

The LHC bunch spacing of 25 ns corresponds to an event rate of 40 MHz. With an average event size of around 1.5 MB, this means that over 60 TB of data is produced every second [67]. Unfortunately, we are not capable of reading out and storing this much data, and have a maximum allowable rate of only ≈1 kHz [68]. It is therefore necessary to reduce the number of events by a factor of 40,000 in a fraction of a second while still retaining events with interesting physics. This is handled by the ATLAS Trigger and Data Acquisition (TDAQ)



(a) Light-jet rejection

(b) Charm-jet rejection

Figure 4.8: Light and charm-jet rejection vs. b-tagging efficiency. The DL1r algorithm is used. DL1 uses the same deep feed-forward neural network structure as DL1r but without the recurrent neural network low-level algorithm. MV2 uses boosted decision trees [66].

system, shown in Figure 4.9.

The TDAQ system consists of 2 levels of triggers. The first level, called the L1 trigger, is hardware-based. Physics events come from one of two subsystems, L1Calo and L1Muon. L1Calo uses information from the calorimeters with reduced granularity. 7200 towers with typical $\Delta\eta \times \Delta\phi$ sizes of 0.1×0.1 feed information into the Cluster Processor (CP) and the Jet/Energy-sum Processor (JEP). The CP identifies electrons, photons, and τ leptons with transverse energy E_{T} above a given threshold in a region of interest, while the JEP identifies jets and the total E_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ [68, 69]. L1Muon uses the information from the RPCs and TGCs to identify high-p_{T} muons. These are augmented by L1Topo, a topological trigger that combines the L1Calo and L1Muon results with geometric information to more accurately calculate event energies. Finally, the information is sent to the Central Trigger Processor (CTP), which combines the information and handles dead-time. The L1 trigger takes $< 2.5\ \mu s$ and brings the event rate down to 100 kHz [68]. The rates for the different trigger types at the L1 level are shown in Figure 4.10.

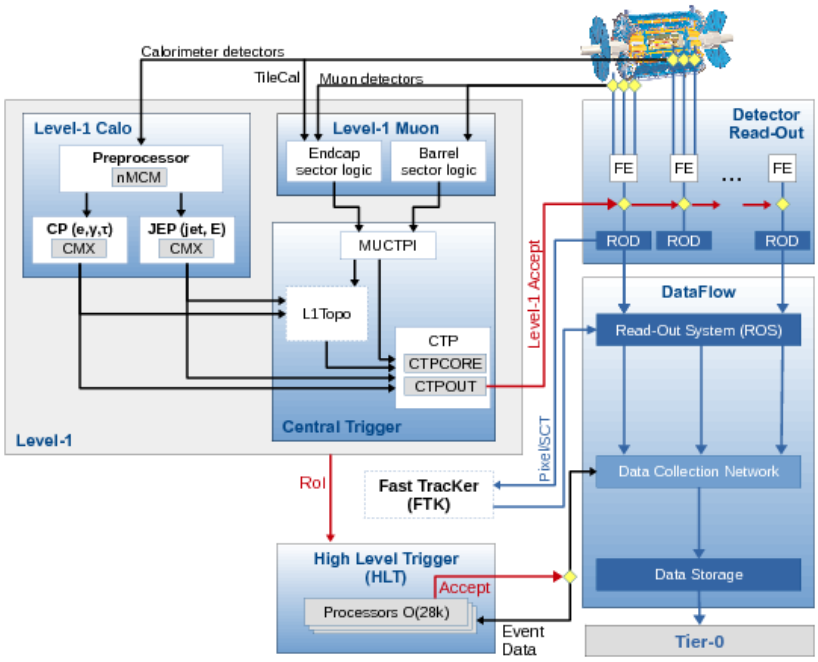


Figure 4.9: Schematic of the ATLAS Trigger and Data Acquisition system [68].

Events that pass the L1 trigger are used as inputs for the High-Level Trigger (HLT). Unlike the L1 trigger, the HLT uses software algorithms rather than custom hardware. The HLT uses higher precision information from the calorimeters and muon spectrometer, as well as tracking information from the inner detector. Most triggers use a 2-pass system. Quick algorithms are used to provide initial event selection, while remaining events are processed using more precise algorithms similar to the offline reconstruction. Some algorithms use information from L1 Regions-of-Interest (ROIs) rather than the full detector to lower the processing time . These are all run on a CPU farm of ≈40,000 Processing Units [70].

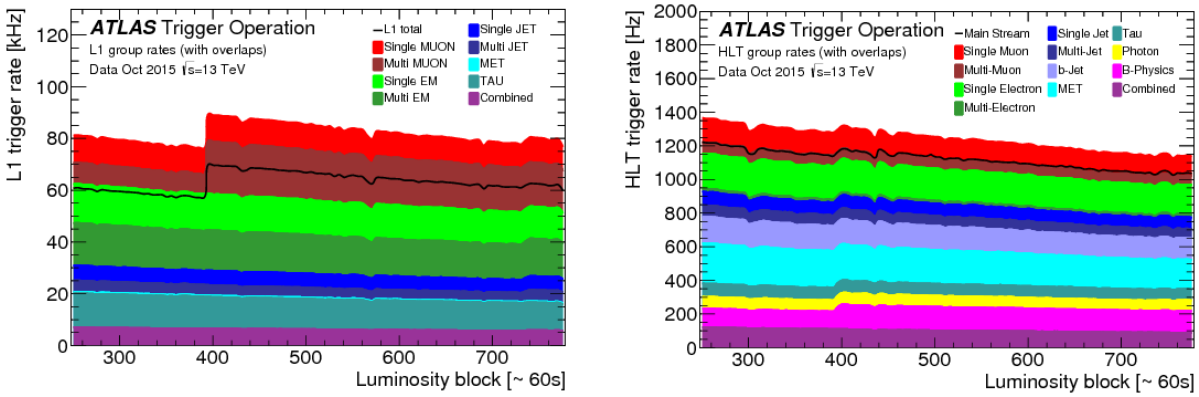
The tracking algorithms are of particular note for this thesis. These algorithms are only run on ROIs around L1 trigger objects. The tracking information is used by the b-jet triggers (among others). These triggers are used to identify jets arising from bottom quark hadronization and parton showers. As described above, b-hadrons travel several millimeters before showering, causing a gap between the primary vertex and the shower start location (called the secondary vertex). Information such as this is used to assign jets a score corre-

sponding to their likelihood of being a b-jet. The triggers then require some number of jets with p_{T} above a threshold and sufficient probabilities of being a b-jet. Many triggers also have requirements on the number and p_{T} of all jets, including those which are not b-tagged. Various triggers target different phase spaces of events by using different values for the number of jets, p_{T}, and probability required [68].

Improved tracking information in the trigger could allow for more efficient b-jet triggers. This was the goal of the Fast TracKer (FTK) [71] and Hardware Tracking for the Trigger (HTT) [72] projects, which were planned to be implemented for Run 3 and the High-Luminosity LHC (HL-LHC) respectively. These projects were designed to use custom firmware and pattern recognition to reduce the time required for tracking, allowing tracks to be provided to the HLT globally rather than in limited ROIs. Due to difficulties in their implementations and improvements in software tracking, these projects were canceled.

The HLT brings the event rate down to the final value of 1.2 kHz. This process takes an average of 235 ms, but a small fraction of events take as long as 8 seconds [68]. The rates for the different categories of trigger at the HLT level are shown in Figure 4.10.

The two stand-alone analyses in this thesis both use a series of b-jet triggers, while the di-Higgs combination analysis uses different triggers for each analysis in the combination. The triggers used for the SUSY analysis are listed in Chapter 7 and their modeling is discussed in detail in Chapter 8. The triggers used for the $hh \to b\bar{b}b\bar{b}$ analysis are discussed in Section 16.3.



(a) L1 (b) HLT

Figure 4.10: Trigger rates for the L1 and HLT triggers [68].

Part III

Search for Supersymmetry

Chapter 5

DATA AND MC SAMPLES

5.1 Data Samples

The ATLAS experiment has a total of 139 fb^{-1} of good quality data for the Run 2 data-taking period, spanning 2015-2018. However, the 2015 data is not used as it corresponds to only 3.2 fb^{-1}. This would add minimal sensitivity to the analysis while requiring separate treatment for trigger calibrations and background modeling. Furthermore, the b-jet triggers used in this analysis require accurate online vertexing and beamspot information. Applying dedicated quality criteria to enforce these requirements reduces the available luminosity to 126 fb^{-1}. The breakdown by year is shown in Table 5.1. Events corrupted due to errors in the Tile, LAr, or SCT detectors are also removed.

5.2 Monte Carlo Samples

In addition to data, Monte Carlo simulations are needed to model signal and background processes. As this analysis uses a purely data-driven background estimation strategy, simulations are only truly needed for the signal processes. However, $t\bar{t}$ samples are also generated to test methods for background reduction as well as to validate the background estimate.

Samples for the signal process are generated using MadGraph v2.6.2 at leading order

Year	Total ATLAS luminosity (fb^{-1})	Luminosity used (fb^{-1})
2015	3.2	-
2016	33.0	24.6
2017	44.3	43.7
2018	58.5	57.7

Table 5.1: The integrated luminosity per year.

(LO) with the NNPDF 2.3 Parton Density Function (PDF) sets [73]. Up to two additional partons are allowed in the simulation. MadGraph v2.6.1 was used for the 130 and 200 GeV mass points. Parton showering and hadronization are simulated using Pythia v8.230 [74]. The cross sections for the processes are calculated at next-to-leading order (NLO) with next-to-leading-log (NLL) precision in the limit of mass-degenerate higgsinos $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, and $\tilde{\chi}_1^\pm$. The cross sections are summed over the $\tilde{\chi}_1^0\tilde{\chi}_2^0$, $\tilde{\chi}_1^0\tilde{\chi}_1^\pm$, $\tilde{\chi}_2^0\tilde{\chi}_1^\pm$, and $\tilde{\chi}_1^\pm\tilde{\chi}_1^\pm$ channels. All the other sparticles assumed to be heavy and decoupled [75–77]. Samples are generated for 130, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, and 1500 GeV higgsino masses, with the gravitino assumed to have negligible mass. The response of the ATLAS detector is simulated using Geant4 [78].

QCD multijet processes dominate the background for this analysis. However, $t\bar{t}$ plays an increasingly important role as $E_{\mathrm{T}}^{\mathrm{miss}}$ increases. This is mainly due to leptonic and semi-leptonic $t\bar{t}$ decays, as the neutrinos produce real $E_{\mathrm{T}}^{\mathrm{miss}}$. As QCD multijet simulations have large theory uncertainties, these backgrounds are not simulated but instead estimated using a data-driven approach. The $t\bar{t}$ samples are generated at NLO with the Powheg Box generator [79–81] at next-to-leading order (NLO) with NNPDF 3.0 PDFs and interfaced to Pythia v8.230 for showering and hadronization. The EvtGen package [82] is used to model hadronic decays, and the h_{damp} parameter in Powheg is set to 1.5 times the mass of the top quark ($m_{\mathrm{top}} = 172.5$ GeV). The predictions of the $t\bar{t}$ MC samples are normalized to the cross sections predicted at next-to-next-to-leading order (NNLO) with NLL precision. As with the signal samples, the response of the ATLAS detector is simulated using Geant4 [78].

Chapter 6

OBJECT DEFINITIONS

This analysis makes use of several types of objects and variables. Four categories (preselected, loose, baseline, and signal) of electrons and muons are defined, while three types (loose, baseline, and signal) of jets are defined.

Preselected electrons are required to pass the loose likelihood identification criteria, have at least one B-layer hit, and have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.47$. Loose electrons are the subset of these with $p_{\mathrm{T}} > 8$ GeV, while baseline electrons are the subset with $p_{\mathrm{T}} > 20$ GeV. Signal electrons are further required to pass overlap removal, loose isolation criteria, and medium identification criteria [83]. Baseline electrons are used in overlap removal. Preselected, loose, and signal electrons are used to construct event vetoes.

Preselected muons are required to pass medium quality requirements, have $p_{\mathrm{T}} > 6$ GeV, and have $|\eta| < 2.7$. Loose muons are the subset of these with $p_{\mathrm{T}} > 8$ GeV, while baseline muons are the subset with $p_{\mathrm{T}} > 20$ GeV. Signal muons are the subset of baseline muons that pass overlap removal, pass tight isolation requirements, and have $|\eta| < 2.5$ [84]. As with electrons, the preselected, loose, and signal muons are used to construct event vetos. Baseline muons are used for overlap removal and are also used to correct the four-momentum of semileptonic decays in b-jets. If any muons fall within $\Delta R < 0.4$ of a jet, the closest such muon is treated as part of the jet.

The preselected and loose lepton categories were defined to allow for orthogonalization with other EWK analyses targeting higgsino pair production. These analyses target different Higgs decay channels, as well as Z boson decays [85–87]. In particular, the Z channels can target leptonic final states. These lepton categories were defined in addition to the baseline category as it was found that using leptons with low p_{T} thresholds for overlap removal caused a loss of signal efficiency of up to 10% for higgsinos, while the scheme discussed in Chapter 7 loses $< 5\%$ of our signal.

Small-radius particle flow jets (R=0.4) are reconstructed using the anti-k_t algorithm.

Baseline jets are required to have $p_{\mathrm{T}} > 20$ GeV. These are used for overlap removal. Signal jets are required to pass overlap removal, have $p_{\mathrm{T}} > 40$ GeV, and have $|\eta| < 2.8$. Loose jets, used for calculating the scalar sum of jet p_{T} (H_{T}), have a lower threshold of $p_{\mathrm{T}} > 25$ GeV. This is done to better match the trigger conditions, as this variable is only used to select which triggers to use. To suppress pile-up jets, loose and signal jets with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$ are required to have a JVT (jet vertex tagger) score greater than 0.2 [88]. This is the medium criteria for the jet vertex tagger. Signal jets used for Higgs candidate construction are required to have $|\eta| < 2.5$ to match the $|\eta|$ requirements for b-jets.

Jets are tagged as containing bottom quarks using the DL1r algorithm at the 77% working point (WP). Of the four possible working points (60%, 70%, 77%, and 85%) defined by ATLAS, this was estimated to yield the highest sensitivity. This working point has rejection factors of 130, 14, and 4.9 for c, τ, and light-flavor jets respectively [64].

An overlap removal procedure is performed using the baseline electrons, muon, and jets. First, any electrons sharing an ID track with a muon are removed. For any remaining electrons, any jets with $\Delta R(j, e) < 0.2$ are removed. Any electrons with $\Delta R(j, e) < \min(0.4, 0.04 + 10\ \mathrm{GeV}/p_{\mathrm{T}}^{e})$ from a remaining jet are removed. Next, muons and jets are considered overlapping if $\Delta R(j, \mu) < 0.2$ or the muon ID track is ghost associated to the jet. Any overlapping jets are removed if they have less than three associated tracks with $p_{\mathrm{T}} > 500$ MeV. Finally, the muons with $\Delta R(j, \mu) < \min(0.4, 0.04 + 10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu})$ to remaining jets are removed.

In addition to the objects, 3 variables must be defined. The missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the magnitude of the negative vector sum of p_{T} of all selected and calibrated physics objects in the event, with an extra term added to account for soft energy in the event that is not associated to any of the selected objects [89]. This soft term is calculated from inner detector tracks matched to the primary vertex to make it more resilient to pile-up contamination. H_{T} is defined as the scalar sum of jet p_{T}. This uses the loose jet collection in order to be more closely match the value calculated in the trigger. The effective

Object	Preselected	Loose	Baseline	Signal
Electrons	$p_{\mathrm{T}} > 7$ GeV, $\|\eta\| < 2.47$, Loose ID	$p_{\mathrm{T}} > 8$ GeV, $\|\eta\| < 2.47$, Loose ID	$p_T > 20$ GeV, $\|\eta\| < 2.47$, Loose ID	Baseline, Medium ID, Loose isolation, pass OR
Muons	$p_{\mathrm{T}} > 6$ GeV, $\|\eta\| < 2.7$, Medium ID	$p_{\mathrm{T}} > 8$ GeV, $\|\eta\| < 2.7$, Medium ID	$p_{\mathrm{T}} > 20$ GeV, $\|\eta\| < 2.7$, Medium ID	Baseline, Tight isolation, $\|\eta\| < 2.5$, pass OR
Jets	-	Baseline, $p_{\mathrm{T}} > 25$ GeV, $\|\eta\| < 2.8$, pass OR, Medium JVT	$p_{\mathrm{T}} > 20$ GeV	Baseline, $p_{\mathrm{T}} > 40$ GeV, $\|\eta\| < 2.8$, pass OR, Medium JVT
b-jets	-	-	Baseline jet, $\|\eta\| < 2.5$, tagged at 77% WP	Signal jet, $\|\eta\| < 2.5$, tagged at 77% WP

Table 6.1: Summary of object definitions. "Baseline" and "Loose" are subject to all the requirements of "Preselected", while "Signal" are subjected to all the requirements of "Preselected" and "Baseline". b-jets of a given category have all the requirements of jets of the same category, plus any listed selections.

mass m_{eff} is defined as the $E_{\mathrm{T}}^{\mathrm{miss}}$ plus the scalar sum of jet p_{T} for the four Higgs candidate jets (defined in Chapter 7).

Chapter 7

EVENT SELECTION

In order to make this analysis as sensitive as possible, we want to minimize the amount of background while still keeping as much of the signal as possible. The background events for this analysis can be split into two categories. The first category is events that can be distinguished from the signal model by their kinematics. These events can be removed via cuts, and will be discussed in this section. The second category of events cannot be easily distinguished from signal events, and thus must be estimated as described in Chapter 9. In addition to removing backgrounds, event selections are required to ensure that all events used are well-understood and properly calibrated.

After data-cleaning, events are required to pass trigger selections. Due to the large quality of data generated by the ATLAS detector, events are required to pass online requirements before being stored for further reconstruction. This analysis uses the logical OR of triggers targeting b-jet final states. In order to properly calibrate these triggers (see Chapter 8), offline kinematic selections are used to sort events into three orthogonal buckets. The buckets are determined using the leading jet p_{T} and H_{T} (sum of scalar loose jet p_{T}) of events, with thresholds optimized for sensitivity for each year. For each bucket, only a single trigger is checked. For 2016, HLT_j100_2j55_bmv2c2060_split is used if the leading jet p_{T} is $>$ 150 GeV. This trigger requires events to have at least one jet with $p_{\mathrm{T}} > 100$ GeV ("j100") plus at least 2 jets with $p_{\mathrm{T}} > 55$ GeV ("2j55"), with the latter 2 b-tagged using the MV2 algorithm at a 60% efficiency working point (WP) ("bmv2c2060"). For all other 2016 events, HLT_2j35_bmv2c2060_split_2j35_L14J15.0ETA25 is used. This requires at least 2 b-tagged jets (60% WP) with $p_{\mathrm{T}} > 35$ GeV plus 2 additional jets with $p_{\mathrm{T}} > 35$ GeV.

For 2017, HLT_j110_gsc150_boffperf_split_2j35_gsc55_bmv2c1070_split_L1J85_3J30 is used if the leading jet p_{T} is $>$ 350 GeV. This requires 1 jet with $p_{\mathrm{T}} > 110$ GeV and 2 b-tagged jets (70% WP) with $p_{\mathrm{T}} > 35$ GeV. Remaining events with $H_{\mathrm{T}} > 850$ GeV use the HLT_2j35_gsc55_bmv2c1050_split_ht300_L1HT190_J15s5.ETA21 trigger, requiring 2 b-tagged

jets (50% WP) with $p_{\mathrm{T}} > 35$ GeV plus $H_{\mathrm{T}} > 300$ GeV. All other 2017 events use HLT_2j15_gsc35_bmv2c1040_split_2j15_gsc35_boffperf_split_L14J15.0ETA25. This requires 2 b-tagged jets (40%) with $p_{\mathrm{T}} > 15$ GeV and 2 additional jets with $p_{\mathrm{T}} > 15$ GeV. This 2b2j trigger has notably lower p_{T} thresholds but less efficient b-tagging thresholds than the corresponding triggers for 2016 and 2018.

For 2018, the HLT_j110_gsc150_boffperf_split_2j45_gsc55_bmv2c1070_split_L1J85_3J30 trigger is used if the leading jet p_{T} is > 500 GeV. The HLT_2j45_gsc55_bmv2c1050_split_ht300_-L1HT190_J15s5.ETA21 is used for remaining events if the $H_{\mathrm{T}} > 700$ GeV. HLT_2j35_-bmv2c1060_split_2j35_L14J15.0ETA25 is used for all other events.

For each year, the first trigger (referred to as the "Initial-State Radiation (ISR) trigger") targets events with $\geq$1 high-p_{T} jet plus $\geq$2 b-tagged jets. The second trigger (called the "H_{T} trigger") targets events with at least two b-tagged jets and high H_{T}. This trigger is not used for the 2016 data-taking period. The third trigger (the "2b2j trigger") captures events with at least two b-tagged jets plus two other jets.

Events are then required to have $\geq$4 b-tagged jets, with an additional category containing exactly two b-tagged jets defined for background estimation purposes. These jets are then used to construct two Higgs boson candidates, requiring four jets. Events are also required to have at least 4 EMTopo jets, though this selection was found to be looser than the particle flow jet requirement. If there are more than four b-jets, the four with the highest p_{T} are selected. For the 2b samples, reconstruction is performed using the two b-jets and two randomly selected non-b-tagged jets.

Given four jets, there are three ways to pair them into Higgs boson candidates. Three methods were considered. In the first method, "min-diff", the pairing is selected to minimize the difference $|m_{h1} - m_{h2}|$. The second method, "min-dhh", selects the pairing that minimizes the planar distance from (m_{h1}, m_{h2}) to (120, 110) GeV. In the final method, "min-dR", the angular distance ΔR_{jj} between the two jets from each Higgs decay is calculated for each possible pairing. $\Delta R_{jj,\mathrm{far}}$ is then defined as the greater of the two angular separations. The

pairing that minimizes this quantity is chosen. The min-diff method ensures that the two Higgs candidates have similar masses, the min-dhh method ensures that they have masses close to the expected values for real Higgs bosons, and the min-dR method ensures that the angular separations between the Higgs boson candidate decay products are small. Real Higgs bosons are expected to meet all three of these conditions.

The min-dR method was used for this analysis. While it has the lowest signal acceptance of the three methods, it was found to provide the strongest sensitivity due to its strong background rejection. Additionally, this method gives a smooth background peaked outside of the analysis regions (see Figure 7.7). This differs from the min-diff and min-dhh methods, where the background is strongly peaked near the line $m_{h1} = m_{h2}$. The background estimation methods described in Chapter 9 are expected to be more reliable for smoother backgrounds such as the one given by the min-dR method.

One of the largest backgrounds for this analysis is $t\bar{t}$, as it is capable of producing numerous b-jets. In particular, semileptonic $t\bar{t}$ is a leading background at high $E_{\mathrm{T}}^{\mathrm{miss}}$ as it produces neutrinos, which are invisible to the ATLAS detector. These events also contain real leptons, which do not appear in the signal model. Therefore, the contribution from $t\bar{t}$ can be reduced by vetoing events containing leptons (here used to mean electrons or muons). Any events with such leptons matching the signal criteria in Chapter 6 are rejected.

We would also like to avoid overlap with other analyses, some of which use lower p_{T} thresholds for their lepton definitions. We therefore veto events which have two or more leptons matching the loose criteria and events which have three or more leptons matching the preselection criteria. These vetoes will make it easier to combine the analysis with others targeting similar higgsino production mechanisms using different final states.

This veto still leaves a substantial $t\bar{t}$ background, especially as it does not affect hadronic $t\bar{t}$ decays. An additional veto was thus developed to target events where one or more of the top quarks decays hadronically. Top quarks nearly always decay as $t \to W^{\pm}b$, and hadronically-decaying $W^{\pm}$ bosons produce two jets. This gives a final state with three jets,

at least one of which must be a b-jet. In addition, one or more of the quarks from the $W^{\pm}$ boson is likely to be an up, down, or strange quark. These form light-flavor jets, which have a low probability to be mis-tagged as a b-jet.

For each event, top candidates are constructed using three jets. One of the Higgs candidate jets is selected as a candidate for the b-jet from the top quark decay. The $W^{\pm}$ boson candidate is then constructed using any two remaining jets, subject to the restriction that at most one of them can be a Higgs candidate jet. This restriction is used because Higgs candidate jets should be heavy-flavor b-jets, while one of the $W^{\pm}$ jets should be light.

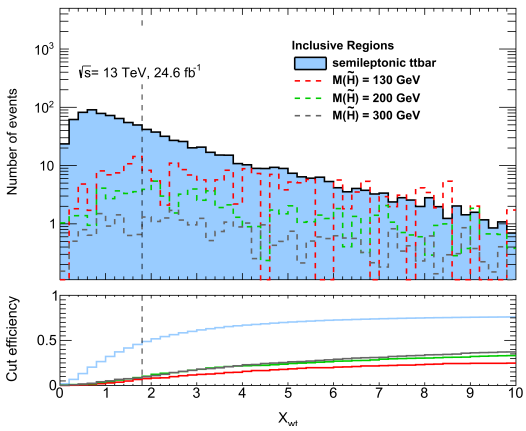
For each possible candidate, the quantity

$$X_{Wt} = \sqrt{\left(\frac{m_W - 80.4 \text{ GeV}}{0.1 m_W}\right)^2 + \left(\frac{m_t - 172.5 \text{ GeV}}{0.1 m_t}\right)^2} \tag{7.1}$$

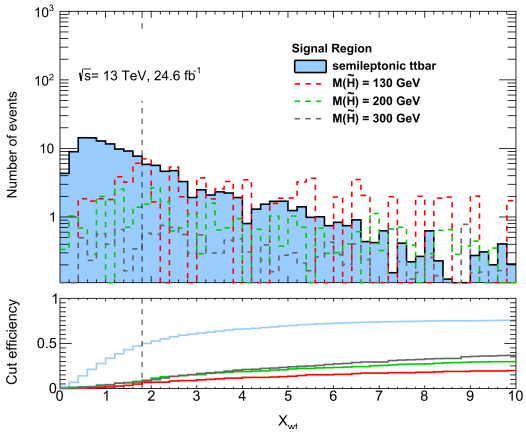
is calculated. For true $t\bar{t}$ events, there should be a combination where the $W^{\pm}$ candidate mass is close to the true value of 80.4 GeV and the top candidate mass is close to the true value of 172.5 GeV, resulting in a small value of X_{Wt}. Events are rejected if $X_{Wt} < 1.8$ for any possible jet combination. The effect of this cut on signal and $t\bar{t}$ is shown in Figures 7.1 and 7.2.

Three types of regions are used in this analysis. The control region (CR) is used to estimate the background and provide systematic uncertainties on the background estimate. The validation region (VR) is used to estimate additional systematic uncertainties and verify that the background estimate is performing reasonably. Finally, the signal region (SR) is used to search for signs of new physics. A well-designed signal region will have a high S/B, while good CRs and VRs will have low signal contamination so as not to bias the background estimate.

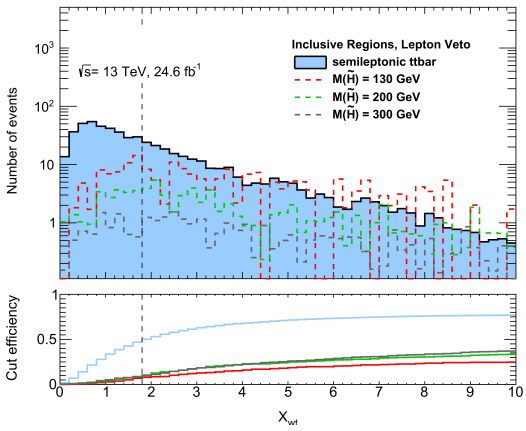
Regions are defined using the masses of the two reconstructed Higgs candidates. The leading (sub-leading) Higgs candidate is defined as the one with greater (lesser) p_{T}. The leading Higgs candidate typically has higher mass. Three variables are defined to determine the regions:



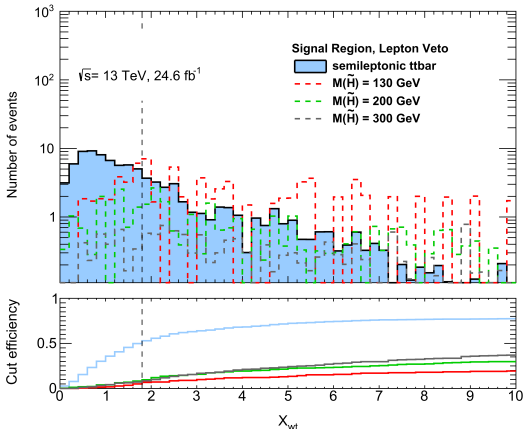
(a) Inclusive, before lepton vetoes



(b) Signal Region, before lepton vetoes

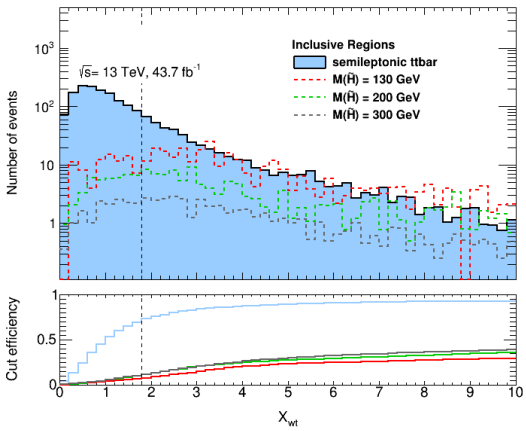


(c) Inclusive, after lepton vetoes

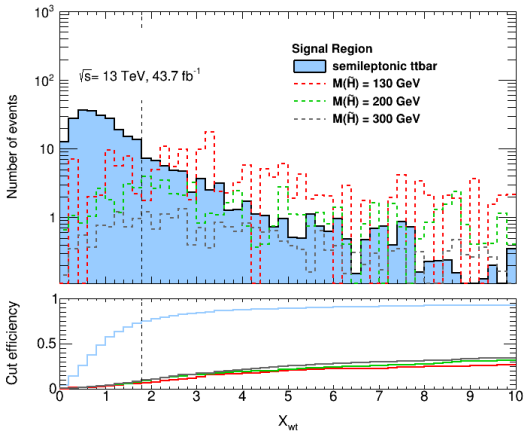


(d) Signal Region, after lepton vetoes

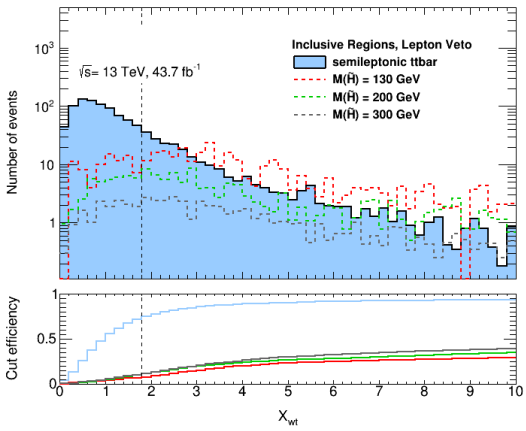
Figure 7.1: X_{Wt} distributions for 130 GeV, 200 GeV, and 300 GeV signals as well as leptonic/ semileptonic $t\bar{t}$ using 2016 simulations.



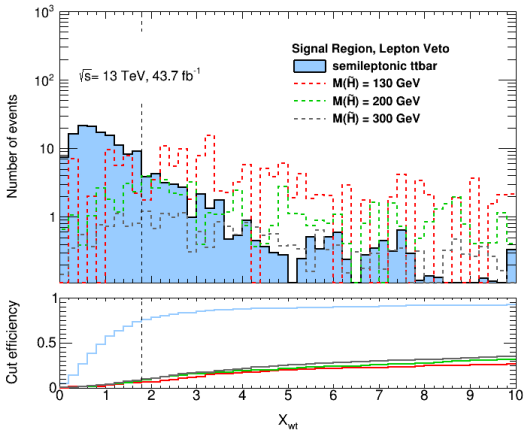
(a) Inclusive, before lepton vetoes



(b) Signal Region, before lepton vetoes



(c) Inclusive, after lepton vetoes



(d) Signal Region, after lepton vetoes

Figure 7.2: X_{Wt} distributions for 130 GeV, 200 GeV, and 300 GeV signals as well as leptonic/ semi-leptonic $t\bar{t}$ using 2017 simulations.

$$X_{hh}^{\mathrm{CR}} = \sqrt{(m_{h,\mathrm{lead}} - 120 * 1.05\ \mathrm{GeV})^2 + (m_{h,\mathrm{subl}} - 110 * 1.05\ \mathrm{GeV})^2} \tag{7.2}$$

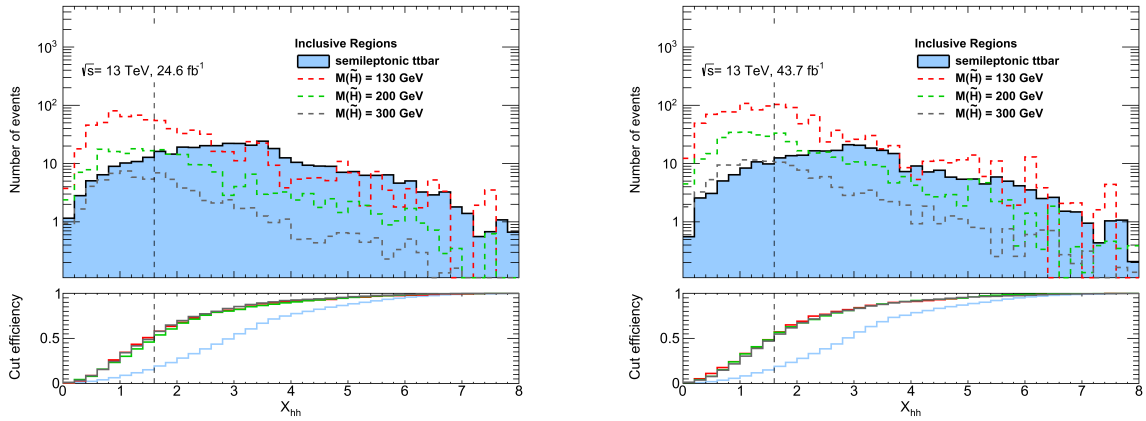
$$X_{hh}^{\mathrm{VR}} = \sqrt{(m_{h,\mathrm{lead}} - 120 * 1.03\ \mathrm{GeV})^2 + (m_{h,\mathrm{subl}} - 110 * 1.03\ \mathrm{GeV})^2} \tag{7.3}$$

$$X_{hh}^{\mathrm{SR}} = \sqrt{\left(\frac{m_{h,\mathrm{lead}} - 120\ \mathrm{GeV}}{0.1 m_{h,\mathrm{lead}}}\right)^2 + \left(\frac{m_{h,\mathrm{subl}} - 110\ \mathrm{GeV}}{0.1 m_{h,\mathrm{subl}}}\right)^2} \tag{7.4}$$

The signal region is defined using X_{hh}^{SR}. The distributions of this variable for 2016 and 2017 are shown in Figure 7.3. An upper threshold of $X_{hh}^{\mathrm{SR}} = 1.6$ is used. The validation region is defined by $X_{hh}^{\mathrm{VR}} < 30$ GeV, with an additional requirement of $X_{hh}^{\mathrm{SR}} \geq 1.6$ to avoid overlap with the signal region. The control region contains all events with $X_{hh}^{\mathrm{CR}} < 55$ GeV, except those that fall into the validation or signal regions.

These regions are selected to have centers near the true Higgs boson mass of 125 GeV, with adjustments for the effects of reconstruction. The denominator terms in the X_{hh}^{SR} definition make it tighter for lower Higgs candidate masses, where the background is larger. Sample signal and background distributions are shown in Figure 7.3. Signal events contain real Higgs bosons, causing the reconstructed masses to be close to the true Higgs mass. This leads to distributions peaked in the signal region. For background events, on the other hand, paired jets do not correspond to true Higgs bosons and the distributions peak outside of the analysis regions. This is desirable as it lowers the background for the analysis.

A summary of the cuts used in this analysis are shown in Table 7.1. The effect of each cut in the analysis is shown in Tables 7.2-7.4. For signal events, the b-jet cut requires 4 or more b-jets. For background events, the b-jet cut requires exactly 2 b-jets to reflect the background estimation strategy. The "good run list" and "run numbers" cuts remove data which lack the online vertexing and beamspot information required to use b-jet triggers. Two candidate events passing the selections are shown in Figures 7.8 and 7.9. The first event has low $E_{\mathrm{T}}^{\mathrm{miss}}$ while the second event has high $E_{\mathrm{T}}^{\mathrm{miss}}$.



(a) 2016 (b) 2017

Figure 7.3: X_{hh}^{SR} distributions for selected signal mass points and leptonic/semileptonic $t\bar{t}$ in the control region. All cuts are made except the SR cut.

2b CR	4b CR	2b VR	4b VR	2b SR	4b SR
Pass b-tagging GRL					
Pass b-jet triggers/buckets					
≥ 4 jets with $p_{\text{T}} > 40$ GeV, $\lvert\eta\rvert < 2.5$					
$== 2$ b-jets	≥ 4 b-jets	$== 2$ b-jets	≥ 4 b-jets	$== 2$ b-jets	≥ 4 b-jets
Signal lepton veto ($p_{\text{T}} > 20$ GeV)					
≤ 1 loose leptons && ≤ 2 preselected leptons					
$X_{Wt} \geq 1.8$ (top veto)					
$X_{hh}^{\text{SR}} \geq 1.6$				$X_{hh}^{\text{SR}} < 1.6$	
$X_{hh}^{\text{VR}} \geq 30$ GeV		$X_{hh}^{\text{VR}} < 30$ GeV		-	
$X_{hh}^{\text{CR}} < 55$ GeV		-			

Table 7.1: A summary of cuts used in the analysis. Lepton and jet definitions are in Table 6.1, while other variable definitions are in Chapter 7. GRL refers to the Good Runs List, which selects events with accurate online vertexing and beamspot information.

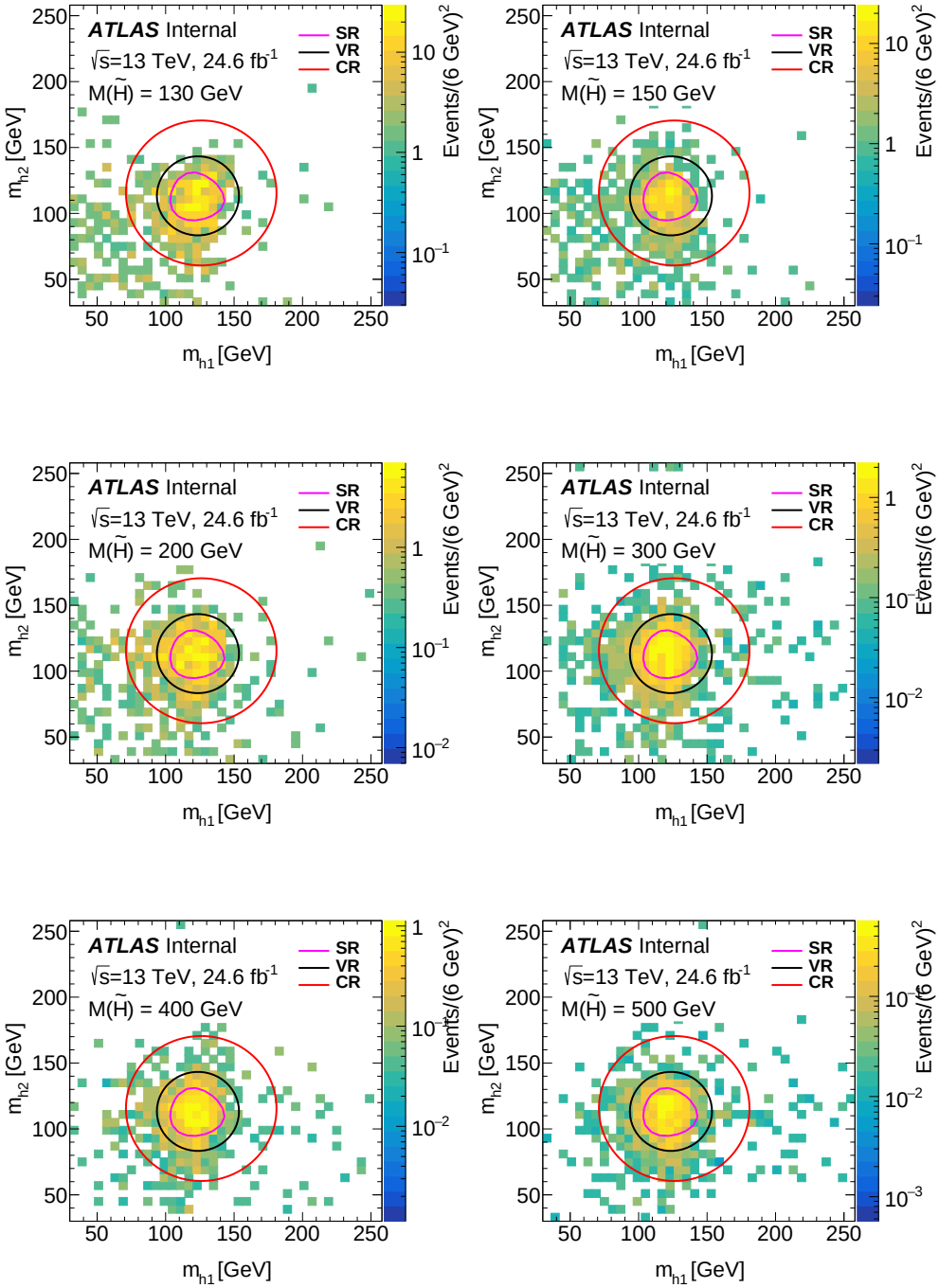


Figure 7.4: Higgs candidate mass distributions for 130, 150, 200, 300, 400, and 500 GeV signals using 2016 simulations. The CR is the region between the red and black ellipses in the Higgs mass plane, the VR is the region between the black ellipse and the pink curve, and the SR is the region within the pink curve.

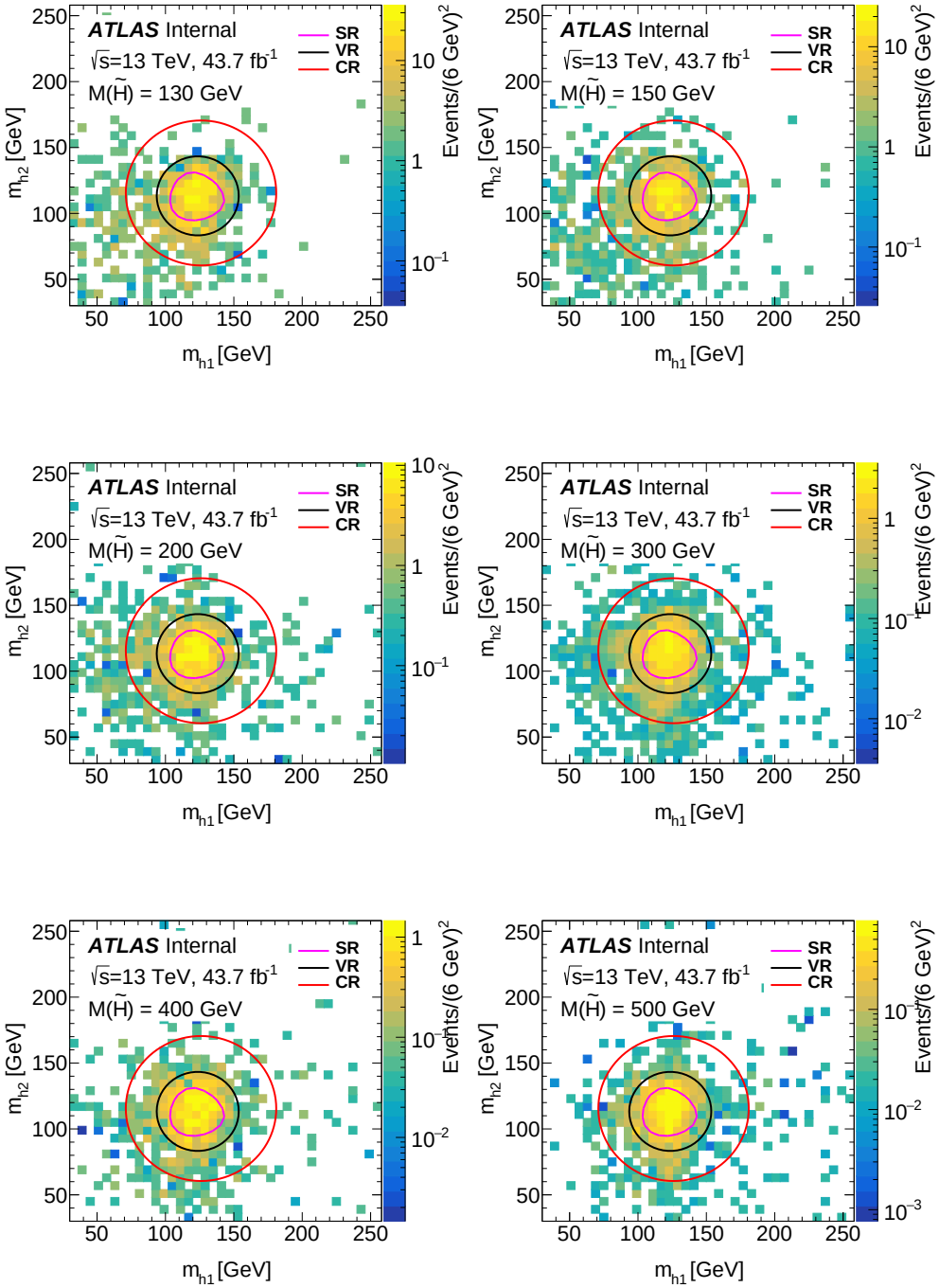


Figure 7.5: Higgs candidate mass distributions for 130, 150, 200, 300, 400, and 500 GeV signals using 2017 simulations. The CR is the region between the red and black ellipses in the Higgs mass plane, the VR is the region between the black ellipse and the pink curve, and the SR is the region within the pink curve.

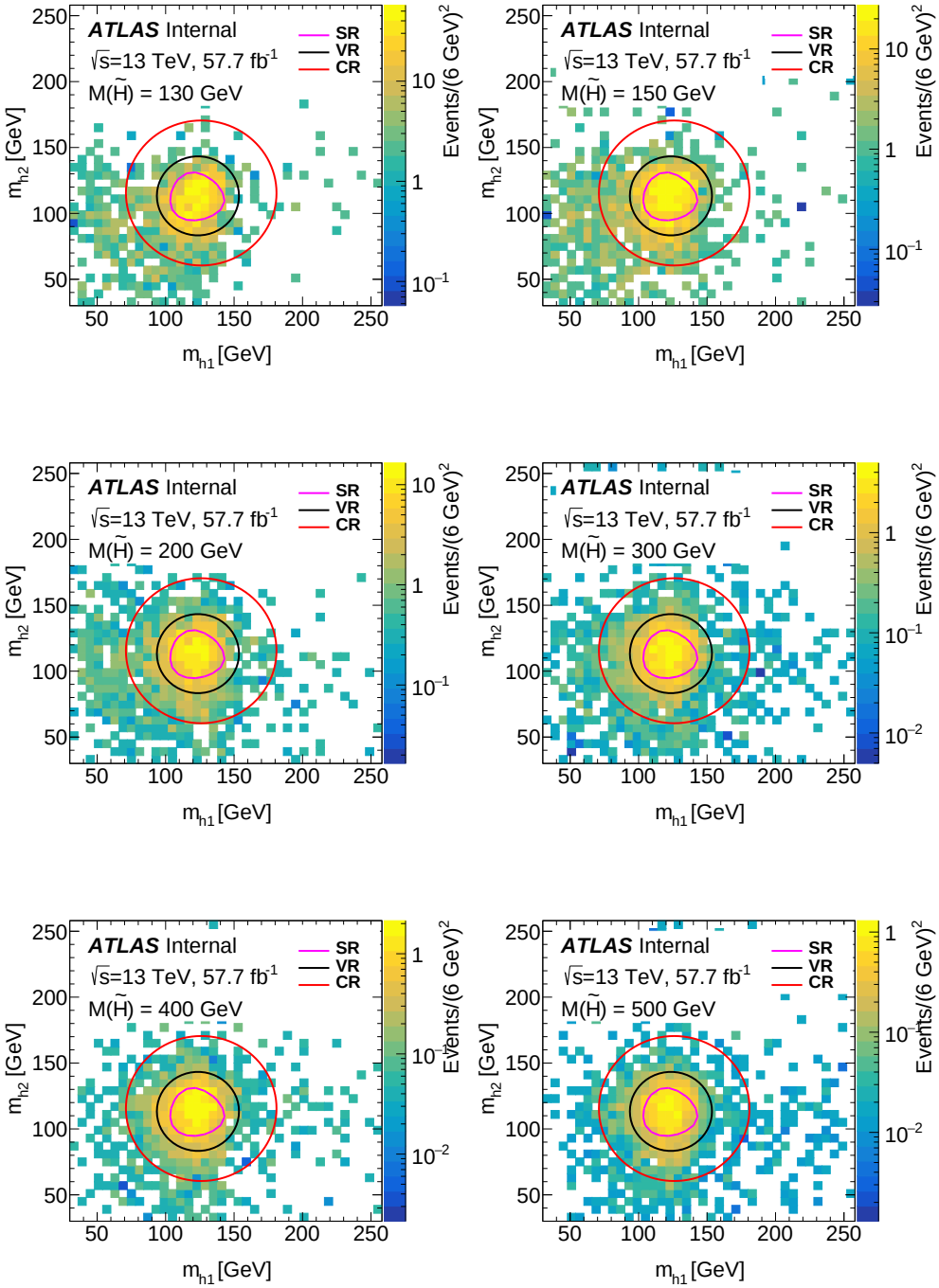


Figure 7.6: Higgs candidate mass distributions for 130, 150, 200, 300, 400, and 500 GeV signals using 2018 simulations. The CR is the region between the red and black ellipses in the Higgs mass plane, the VR is the region between the black ellipse and the pink curve, and the SR is the region within the pink curve.

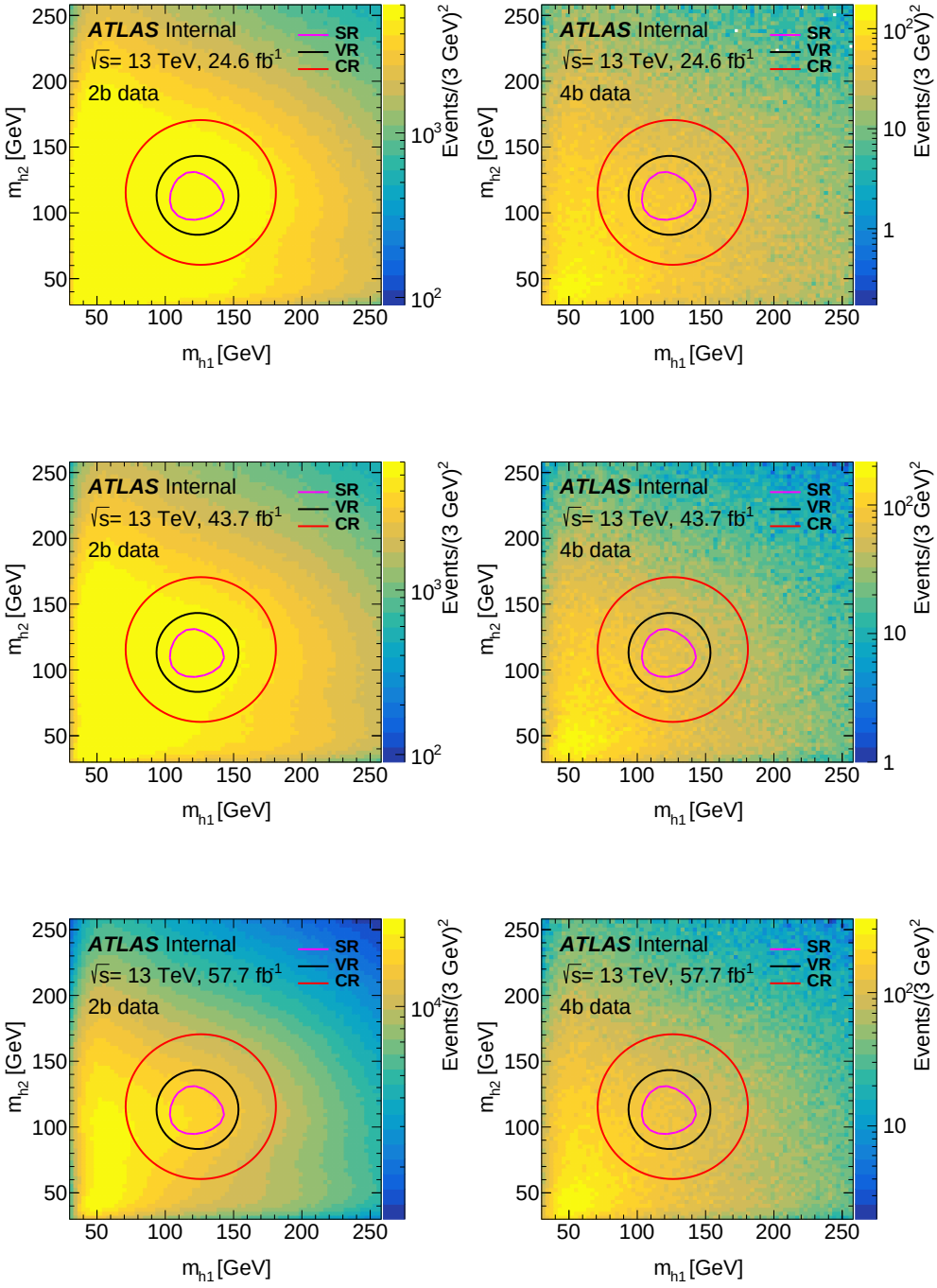


Figure 7.7: Higgs candidate mass distributions for the 2b (left) and 4b (right) samples. The top, middle, and bottom rows show the 2016, 2017, and 2018 distributions respectively. The CR is the region between the red and black ellipses in the Higgs mass plane, the VR is the region between the black ellipse and the pink curve, and the SR is the region within the pink curve.

	Unweighted	Unweighted %	Weighted	Weighted %
Total number of events	1944000	100	532691	100
Events after preselection	111611	5.7413	110030	20.6555
hh decays	51599	2.6543	49946.6	9.3763
Good run list	46958	2.4155	45028.6	8.4530
Run numbers	46958	2.4155	45028.6	8.4530
Trigger	28388	1.4603	25388.5	4.7661
b-jets	3729	0.1918	3491.78	0.6555
Lepton veto	3560	0.1831	3330.97	0.6253
Top veto	3250	0.1672	3048.00	0.5722
Signal Region ($X_{hh}^{\mathrm{SR}} < 1.6$)	1210	0.0622	1183.2	0.2221

Table 7.2: 150 GeV signal MC cutflow. The preselection requires 4 or more jets with $p_{\mathrm{T}} > 25$ GeV and 2 or more b-jets with $p_{\mathrm{T}} > 25$ GeV. The b-jet cut requires exactly 2 b-jets with $p_{\mathrm{T}} > 40$ GeV.

	Unweighted	Unweighted %	Weighted	Weighted %
Total number of events	1944000	100	39594.8	100
Events after preselection	323757	16.6542	25088.8	63.3639
hh decays	140716	7.2385	10873.5	27.4619
Good run list	127883	6.5783	9845.79	24.8664
Run numbers	127883	6.5783	9845.79	24.8664
Trigger	53184	2.7358	3890.06	9.8247
b-jets	7668	0.3944	550.332	1.3899
Lepton veto	7235	0.3722	519.317	1.3116
Top veto	6380	0.3282	459.686	1.1610
Signal Region ($X_{hh}^{\mathrm{SR}} < 1.6$)	2515	0.1294	184	0.4651

Table 7.3: 300 GeV signal MC cutflow. The preselection requires 4 or more jets with $p_{\mathrm{T}} > 25$ GeV and 2 or more b-jets with $p_{\mathrm{T}} > 25$ GeV. The b-jet cut requires exactly 2 b-jets with $p_{\mathrm{T}} > 40$ GeV.

	Unweighted	Unweighted %
Total number of events	431869001	100.0
Good run list	405311292	93.8505
Run numbers	405311292	93.8505
Trigger	222345231	51.4844
b-jets	179023429	41.4532
Lepton veto	175069620	40.5377
Top veto	164536349	38.0987
Signal Region	3854553	0.8925

Table 7.4: Data cutflow for the background estimate. The preselection requires 4 or more jets with $p_{\mathrm{T}} > 25$ GeV and 2 or more b-jets with $p_{\mathrm{T}} > 25$ GeV. The b-jet cut requires exactly 2 b-jets with $p_{\mathrm{T}} > 40$ GeV.

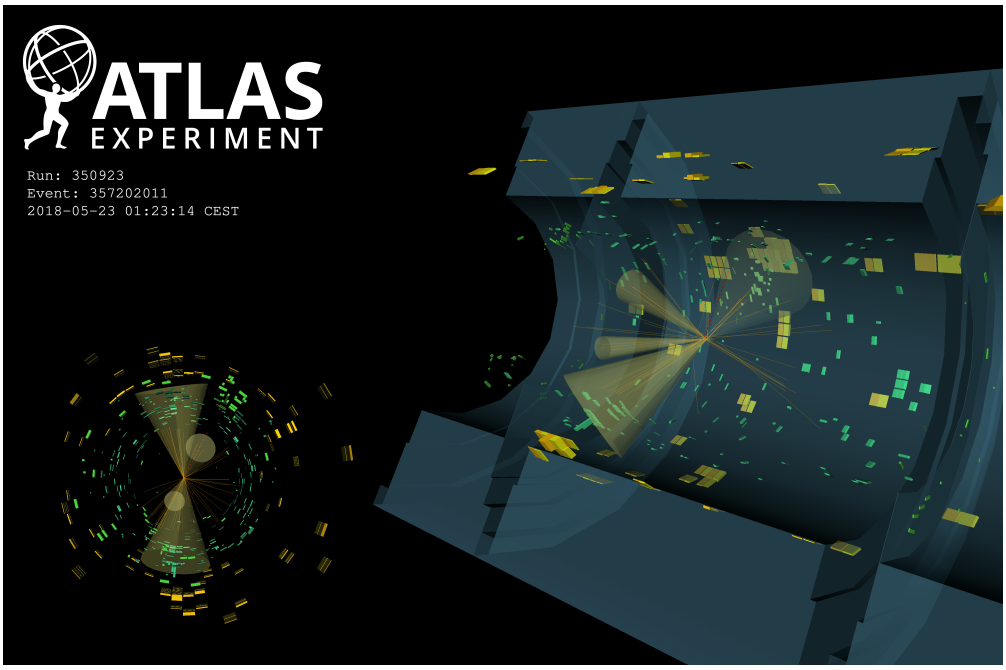


Figure 7.8: An event display for the analysis. The yellow cones are jets, the dashed red line is $E_{\mathrm{T}}^{\mathrm{miss}}$ and the orange lines are tracks. This event has $E_{\mathrm{T}}^{\mathrm{miss}} = 14.0$ GeV and $m_{\mathrm{eff}} = 283.3$ GeV.

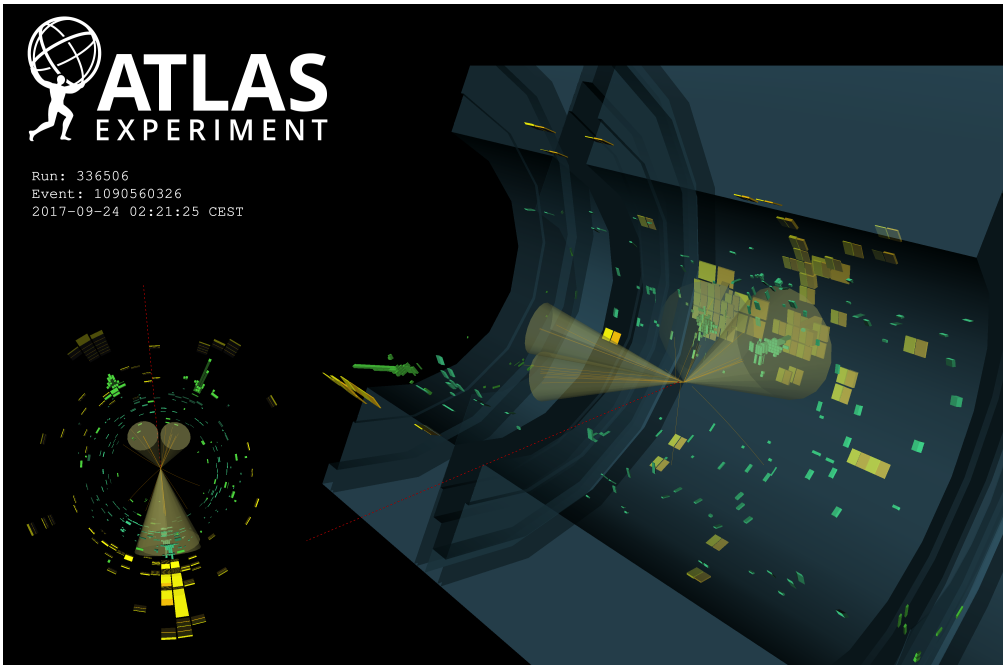


Figure 7.9: An event display for the analysis. The yellow cones are jets, the dashed red line is $E_{\mathrm{T}}^{\mathrm{miss}}$ and the orange lines are tracks. This event has $E_{\mathrm{T}}^{\mathrm{miss}} = 254.7$ GeV and $m_{\mathrm{eff}} = 881.3$ GeV.

Chapter 8

TRIGGER MODELING

Many SUSY analyses make use of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, which captures events with large missing transverse momentum. The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is only fully efficient for $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV (see Figure 8.1) [90]. Events with low-mass higgsinos, however, typically only have $E_{\mathrm{T}}^{\mathrm{miss}} \approx 100$ GeV. The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger would thus reject most of our signal events, which would severely limit the sensitivity of this analysis. To avoid this problem, a combination of b-jet triggers is used.

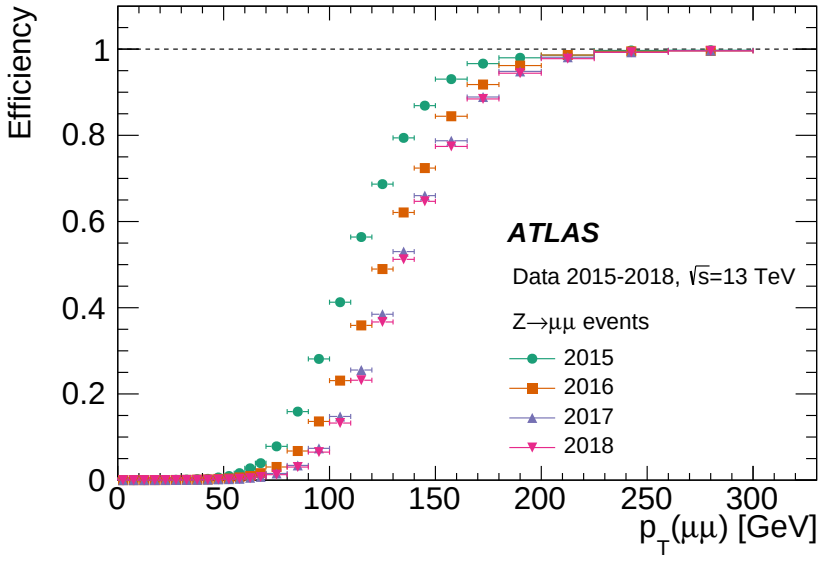


Figure 8.1: Trigger efficiency for the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger using $Z \to \mu\mu$ events [90].

As discussed in Chapter 5, b-jet triggers require accurate online-vertexing and beamspot information. This lowers the available amount of data by about 10%, leaving 126 fb^{-1} for the analysis. We use up to 3 triggers for each year, as this allows us to retain more events than if any single trigger was used. These triggers fall into 3 types. The first type targets initial-state radiation (ISR), selecting events with at least 2 b-jets plus 1 high-p_{T} jet. The second type (referred to hereafter as 2bHT) selects events with at least 2 b-jets plus a scalar sum of jet p_{T} (H_{T}) above a given threshold. The final trigger (2b2j) selects events with at

least 2 b-jets plus 2 or more other jets. Triggers requiring > 2 b-jets were avoided as they would perform differently on our 2b and 4b samples. The H_{T} trigger was not used for 2016 as it provided negligible additional signal yet would require a different calibration scheme than the 2017 and 2018 2bHT triggers. This is because the 2017 and 2018 triggers have an L1 H_{T} requirement, while the 2016 trigger only has an HLT H_{T} requirement.

8.1 *B*-Jet Trigger Scale Factors

ATLAS simulations do not perfectly model data. One of the areas in which these differ is the b-tagging efficiency. This holds both for the offline tagging used in reconstructed samples and the online tagging used for the HLT triggers. This issue is mitigated by applying event-level weights to MC events. These scale factors are defined by

$$w = \frac{\epsilon_{\mathrm{data}}}{\epsilon_{\mathrm{MC}}} \tag{8.1}$$

where w is the event weight and ϵ represents the efficiency of the b-tagging algorithms. A single scale factor must be used for both online and offline efficiencies in order to account for the correlations between them. The ATLAS collaboration uses centrally derived efficiencies which are calulated using $t\bar{t}$ and provided on a per-jet basis [42]. The efficiencies for a given jet depend on the b-tagging working point being used as well as the p_{T} of the jet. Three types of efficiences are provided: online, offline, and conditional (online given offline). These can be combined using Bayesian statistics. The probability of a jet being tagged online and offline is given by

$$\epsilon(on \bigcap off) = \epsilon(on|off)\epsilon(off) \tag{8.2}$$

where ϵ can refer to either the data or MC efficiency. The probability of failing online tagging but passing offline is given by

$$\epsilon(\bar{on} \bigcap off) = [1 - \epsilon(on|off)]\epsilon(off) \tag{8.3}$$

and the probability of passing online tagging but failing offline is given by

$$\epsilon(on \bigcap \bar{off}) = \epsilon(on) - \epsilon(on|off)\epsilon(off) \tag{8.4}$$

Finally, the probability of failing both online and offline tagging is given by

$$\epsilon(\bar{on} \bigcap \bar{off}) = 1 - \epsilon(off) - \epsilon(on) + \epsilon(on|off)\epsilon(off) \tag{8.5}$$

Using these efficiencies requires us to identify which offline jets correspond to the online HLT jets. This is done by matching jets based on their ΔR angular separation. The process starts with the highest p_{T} offline jet. The closest HLT jet is matched to the offline jet, provided their ΔR angular separation is less than 0.4. Otherwise, no HLT jet is matched to the offline jet. This procedure is then repeated for each offline jet in order of decreasing p_{T}. Each HLT jet is only allowed to be matched to one offline jet. For each trigger, events are removed if there are too few offline jets that are both matched to HLT jets and have sufficient b-tagging scores to pass the trigger. This occurs for less than 1% of events.

The total efficiency for each event is given by the product of efficiencies for each jet in the event. The scale factor is then given by Equation 8.1. This process is run for every systematic variation (see Chapter 11) that affects b-tagging. This includes the jet energy scale (JES); jet energy resolution (JER); jet mass scale (JMS); and bottom, charm, and light jet uncertainties.

8.2 Jet and H_{T} Trigger Scale Factors

The triggers used in this analysis have minimum thresholds for b-tagging score, jet p_{T}, and H_{T}. The scale factors due to b-tagging are described above, leaving those due to jet p_{T} and H_{T}.

Scale factors for the jet and H_{T} requirements of the triggers were derived for each trigger by comparing data and MC using a $t\bar{t}$-enriched sample obtained from a single-muon trigger. Jet scale factors are defined for each jet required by the trigger and are given as a function of jet p_{T}, while H_{T} scale factors are derived for the overall event as a function of H_{T}. Separate scale factors are provided for the L1 and HLT portions of the triggers.

The systematic uncertainties of the scale factors are estimated by comparing the efficiencies calculated in the $t\bar{t}$ MC sample and a Standard Model $hh \to b\bar{b}b\bar{b}$ MC sample. The uncertainty is of the order of 10%.

8.3 Trigger Buckets

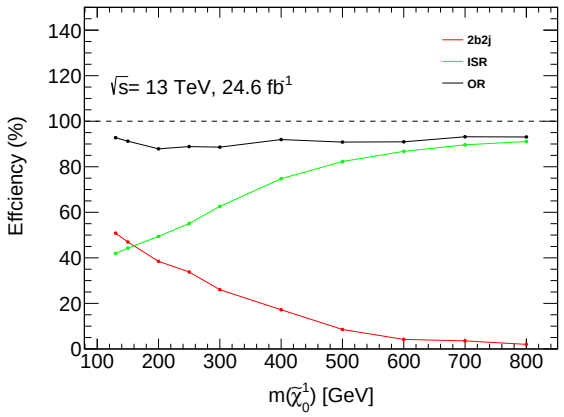
Different triggers have different online b-tagging efficiencies and thus different trigger scale factors. If an event only passes one trigger, it is clear which scale factor to apply. If an event passes more than one trigger, however, we need to select which trigger scale factor to apply. This problem is avoided by sorting events into orthogonal buckets. For 2016 events, a threshold is set on the p_{T} of the leading jet. If an events passes this threshold, it is categorized as an ISR event and the ISR trigger and its scale factor are used. Otherwise, the event is a 2b2j event and the 2b2j trigger and scale factors are used. 2017 and 2018 events are divided similarly, with an additional category of H_{T} events where the p_{T} criteria is not met but the H_{T} of the event is above a threshold value. Events in this category use the H_{T} trigger and scale factors.

This strategy means that for each event, only one trigger is considered. The ISR trigger

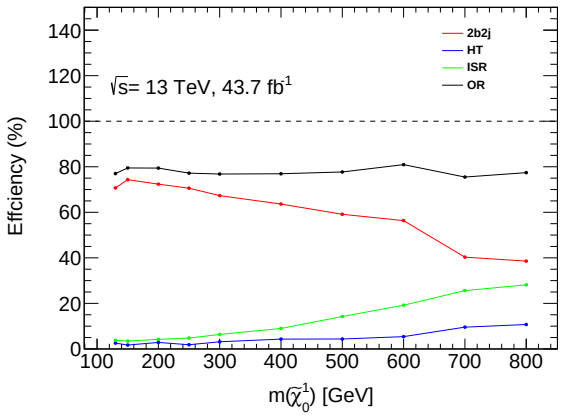
	p_{T} threshold	H_{T} threshold
2016	150 GeV	N/A
2017	350 GeV	850 GeV
2018	500 GeV	700 GeV

Table 8.1: Thresholds for the trigger buckets.

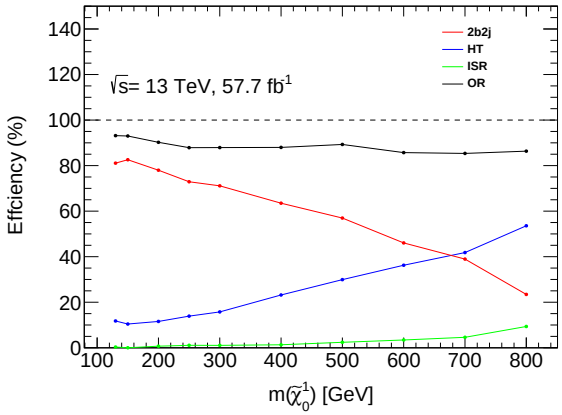
is used for ISR events, the H_{T} trigger is used for H_{T} events, and the 2b2j trigger is used for 2b2j events. While this causes a loss in signal efficiency when compared to accepting events passing any of the three triggers, it also lowers the number of background events passing the triggers. The overall sensitivity remains largely unaffected. The p_{T} and H_{T} thresholds for each year are selected using their resulting sensitivities, with the restriction that each of the three buckets must maintain a sizable number of events. The resulting thresholds are shown in Table 8.1, and the corresponding triggers are listed in Chapter 7. The efficiencies for each bucket and their combination, as measured with signal MC after all selections, are shown in Figure 8.2. Efficiencies of around 90% are found for all target higgsino masses for the 2016 and 2018 data-taking periods. The 2017 signals exhibit a slightly lower efficiency due to different trigger thresholds. Note that the efficiency is not with respect to no trigger requirements, but rather with respect to requiring the logical OR of a large set of lepton, photon, jet, b-jet, and $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers.



(a) 2016



(b) 2017



(c) 2018

Figure 8.2: Trigger efficiencies for each bucket and their logical OR in the SR after all other selections are made. The efficiencies are expressed as the ratio of the number of events passing the trigger and selections to the number of events passing the selections.

Chapter 9

BACKGROUND MODELING

After the selections described in Chapter 7, there are two major backgrounds for this analysis. The largest, QCD multijet, dominates for events with low $E_{\mathrm{T}}^{\mathrm{miss}}$. The sub-leading background, $t\bar{t}$, becomes dominant at high $E_{\mathrm{T}}^{\mathrm{miss}}$ as the neutrinos from leptonic and semi-leptonic top quark decays create real $E_{\mathrm{T}}^{\mathrm{miss}}$.

Monte Carlo modeling of QCD multijet events is relatively poor, as well as computationally expensive. This problem is circumvented through the use of a purely data-driven background estimate. This method, called the ABCD method, models regions with four or more b-tagged jets (called the 4-tag or 4b sample) using regions containing two b-tagged jets (called the 2-tag or 2b sample). We begin by choosing two variables. The number of b-tagged jets and the region (control, validation, or signal) were selected for this analysis. We then define three regions offset from the search region (the 4b SR) by altering each of these selections, as shown in Figure 9.1. This results in a doubly-offset region A (2b CR) and two singly-offset regions B (4b CR) and C (2b SR). All other analysis selections are the same for each region.

Each of the regions A, B, and C has low signal contamination but is close enough to the target region D to have similar kinematics. The fundamental assumption of the ABCD method is that the ratio between events and distributions from A to B is the same as that from C to the background of D. The control region is used rather than the validation region (despite the VR being closer to the target SR) so that the latter can be used to validate the procedure and estimate the systematic uncertainties.

9.1 Normalization

Following the ABCD method, a normalization factor

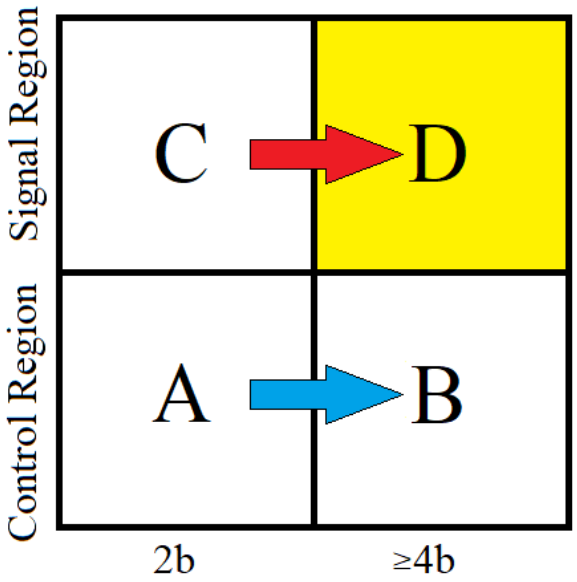


Figure 9.1: Diagram of the ABCD method.

$$\mu_{\mathrm{CR}} = \frac{N_{\mathrm{CR}}^{4b}}{N_{\mathrm{CR}}^{2b}} \tag{9.1}$$

is defined for each year, where N is the number of events in a given region. These are measured to be

$$\mu_{\mathrm{CR}}^{2016} = (7.27 \pm 0.04) * 10^{-3} \tag{9.2}$$

$$\mu_{\mathrm{CR}}^{2017} = (1.208 \pm 0.006) * 10^{-2} \tag{9.3}$$

$$\mu_{\mathrm{CR}}^{2018} = (7.98 \pm 0.03) * 10^{-3} \tag{9.4}$$

where the uncertainties are statistical. The differences between years are due to differing trigger conditions and thresholds.

The 2b data samples are multiplied by these scale factors to get the 4b background estimates. However, as discussed below, the final normalization can be slightly altered by the kinematic reweighting.

9.2 Kinematic Reweighting

While the ABCD method gives a prediction for the number of background events, it does not predict the kinematic distributions of these events. The 2b and 4b kinematics differ due to factors such as uneven contributions from different QCD processes. The final statistical analysis bins the data in $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff}, making it essential that these kinematics are accurately modeled. This is accomplished by reweighting the kinematics of the 2b samples to match those of the 4b samples using boosted decision trees (BDT).

For each year of data, a BDT is trained to reweight the 2b control region to match the 4b control region using a large list of input variables. Using the hep_ml toolkit implementation [91], we define the metric

$$\chi^2 = \sum_{\text{leaf}} \frac{(N^{2b}_{\text{CR,leaf}} - N^{4b}_{\text{CR,leaf}})^2}{N^{2b}_{\text{CR,leaf}} + N^{4b}_{\text{CR,leaf}}} \tag{9.5}$$

Each node splits the dataset into two bins (or "leaves") by selecting the input variable and cut value that maximizes this χ^2. This procedure repeats on the resulting bins until there are 5 layers or further splits would result in < 250 events per node. The resulting bins constitute the first tree. Each 2b CR event is then given a weight

$$e^{\lambda N^{4b}_{\text{CR,leaf}}/N^{2b}_{\text{CR,leaf}}} \tag{9.6}$$

where λ is a parameter called the learning rate and leaf refers to the bin in which the event falls. The weighted events output from the tree are then fed into a new tree. This process is repeated until the desired number of trees is reached. The final weight for each event is the product of the event's weights from each tree. After the procedure is trained in the control region, weights are predicted for all 2b events. This follows the same general principle as the ABCD method. An offset region (the CR) is used to estimate the 4b/2b ratio μ and train a BDT to learn the differences in kinematics between the 2b and 4b regions. This scale factor and BDT are then applied to the 2b SR to estimate the number of events and the kinematics

of the background in the 4b SR.

The weights are normalized such that the sum of weights in the control region is unchanged by the reweighting. The factor derived in this normalization is applied to all weights, including those outside of the control region. It is important to note that this procedure is able to change the total predicted yields in the validation and signal regions. This allows the BDT to learn changes in yields across the Higgs candidate mass plane. The post-reweighting yields in the validation and signal regions are equal to the pre-reweighting yields scaled by the average weights $\bar{w}_{\mathrm{VR}}/\bar{w}_{\mathrm{CR}}$ and $\bar{w}_{\mathrm{SR}}/\bar{w}_{\mathrm{CR}}$ respectively. An alternative procedure where the weights were normalized separately in each region (meaning the post-reweighting yields are determined entirely by the scale factor μ) was also tested. The former method was selected as it more accurately predicted the number of 4b events in the validation region.

As the BDT is able to select the training variables with the best performance, it was provided a large input list of 51 reweighting variables. These are:

- The number of jets in the event
- $E_{\mathrm{T}}^{\mathrm{miss}}$
- The mass, energy, p_{T}, η, and ϕ of each Higgs candidate
- The mass, energy, p_{T}, η, and ϕ of each Higgs candidate jet
- The mass and p_{T} of the di-Higgs system
- The number of track-jets associated with each Higgs candidate
- The average η of the Higgs candidate jets
- The ΔR separation of the jets in each Higgs candidate
- The smaller of the two ΔR separations of the jets in each Higgs candidate
- The larger of the two ΔR separations of the jets in each Higgs candidate

- The sum and difference of the two ΔR separations of the jets in each Higgs candidate
- The ΔR separation between the two Higgs candidates
- X_{Wt} from Equation 7.1, removing the restriction that at least one of $W^{\pm}$ candidate jets must not be a Higgs candidate jet so that the variable is well-defined for all events
- The $\Delta\eta$ separation between the two Higgs candidates
- $|cos(\theta_1)|$, the magnitude of the cosine of the angle between the leading jet of the leading Higgs candidate and the overall leading Higgs candidate
- $|cos(\theta_2)|$, the magnitude of the cosine of the angle between the leading jet of the sub-leading Higgs candidate and the overall sub-leading Higgs candidate
- $|cos(\theta^*)|$, the magnitude of the cosine of the angle between the leading Higgs candidate and the beam axis
- ϕ, the magnitude of the angle between the cross product of the jet vectors in the leading Higgs candidate (in the hh center-of-mass frame) and the negative cross product of the jet vectors in the sub-leading Higgs candidate (in the hh center-of-mass frame)
- ϕ_1, the magnitude of the angle between the cross product of the jets in the leading Higgs candidate (in the hh center-of-mass frame) and the cross product of the beam axis with the leading Higgs candidate vector (in the hh center-of-mass frame)

The BDT features five hyperparameters. The learning rate λ appears in Equation 9.6 and determines how quickly the BDT learns differences between the samples. The minimum number of events per node stops the splitting procedure if the statistics in a node fall below the threshold. The number of layers determines how many node splittings are allowed, while the number of trees sets the number of times the weighting procedure is iterated. Finally, the sampling fraction is the fraction of events used for each tree. A smaller sampling fraction can

Maximum number of layers	4, 5, 6
Minimum number of events per node	125, 250, 500
Number of trees	30, 50, 70, 100
Learning rate	0.1, 0.2, ..., 0.9, 1.0
Sampling fraction	0.4, 0.7

Table 9.1: Hyperparameters tested for the reweighting boosted decision trees.

mitigate overtraining. A scan was performed over various possible hyperparameter values to select those which yielded the best performance. The values tested are shown in Table 9.1.

This yields a total of 720 different combinations. Each of these was tested using the 2017 dataset and applied to the 2b data. A second BDT was then trained as a classifier to distinguish the resulting 4b and reweighted 2b samples. This BDT used all of the reweighting variables from the first BDT, as well as the m_{eff}. The two BDTs are adversarial. The better the first (reweighting) BDT performs, the more similar the 4b and reweighted 2b samples will be. This means that the second (classifier) BDT will have a more difficult time distinguishing the samples, and thus have a worse performance. This is measured using the area under the ROC curve (AUC) of the classifier BDT. An AUC of 1.0 corresponds to perfect classifier performance (poor reweighting performance) while an AUC of 0.5 corresponds to minimal classifier performance (unable to distinguish 4b and reweighted 2b, perfect reweighting performance). Therefore, the best hyperparameters are those which yield an AUC close to 0.5 for the classifier BDT.

As shown in Figure 9.2, performance generally improves with increasing learning rate, maximum layers, and number of trees. This stops once the performance saturates at an AUC of ≈ 0.5. The minimum number of events per node and sampling fraction were found to have minimal impact on the AUC. This scan yielded a large phase space of hyperparameters consistent with maximal reweighting performance. The hyperparameters were selected to

Maximum number of layers	5
Minimum number of events per node	250
Number of trees	50 (2016), 75 (2017), 100 (2018)
Learning rate	0.3
Sampling fraction	0.4

Table 9.2: Hyperparameters selected for the reweighting boosted decision trees.

fall within this phase space while keeping the learning rate and number of trees as small as possible to avoid overtraining. The selected values are shown in Table 9.2

Fewer trees were chosen for years with less data to prevent the BDT from learning the Poisson weights used for bootstrap (Chapter 11) variations.

After training on the control region, the BDT is applied to the full 2b sample. Each event gets a weight, with the weights normalized such that

$$N_{\mathrm{CR}}^{2b,\mathrm{rw}} \equiv \sum_{2b\ \mathrm{CR}} w = \mu_{\mathrm{CR}} * N_{\mathrm{CR}}^{2b} \tag{9.7}$$

for each year of data, where w are the event weights. The reweighted counts in the validation and signal regions are allowed to differ from $\mu_{\mathrm{CR}} * N_{\mathrm{VR(SR)}}^{2b}$, yielding values of

$$N_{\mathrm{VR}}^{2b,\mathrm{rw}} \equiv \sum_{2b\ \mathrm{VR}} w = \bar{w}_{\mathrm{VR}} * \mu_{\mathrm{CR}} * N_{\mathrm{VR}}^{2b} \tag{9.8}$$

$$N_{\mathrm{SR}}^{2b,\mathrm{rw}} \equiv \sum_{2b\ \mathrm{SR}} w = \bar{w}_{\mathrm{SR}} * \mu_{\mathrm{CR}} * N_{\mathrm{SR}}^{2b} \tag{9.9}$$

where $\bar{w}_X$ are the average weights in region X after normalization.

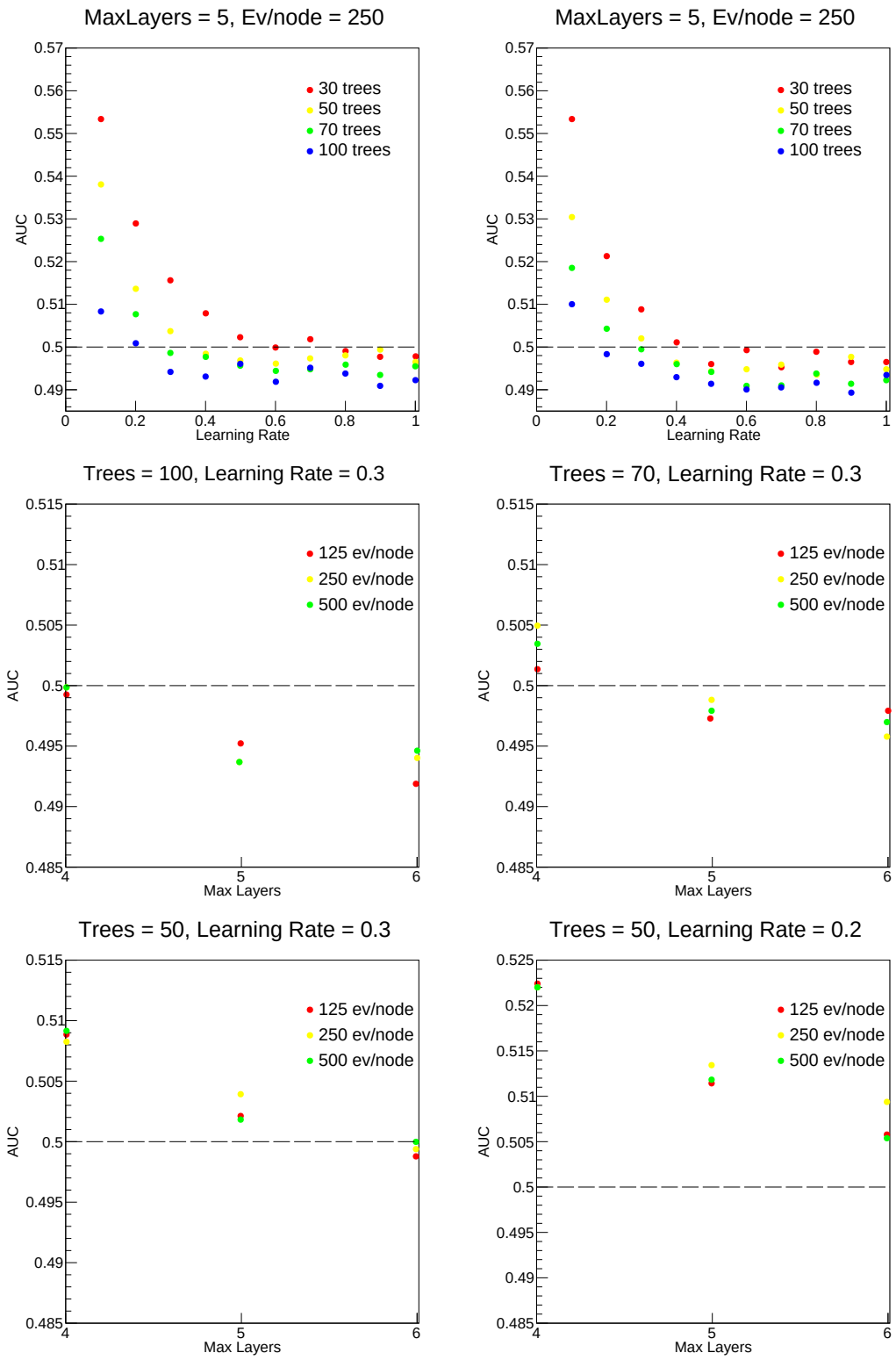


Figure 9.2: AUCs for classifier BDTs using 2017 data. The sampling fraction is 0.4 for all points shown except in the upper right plot, which shows a sampling fraction of 0.7 for comparison.

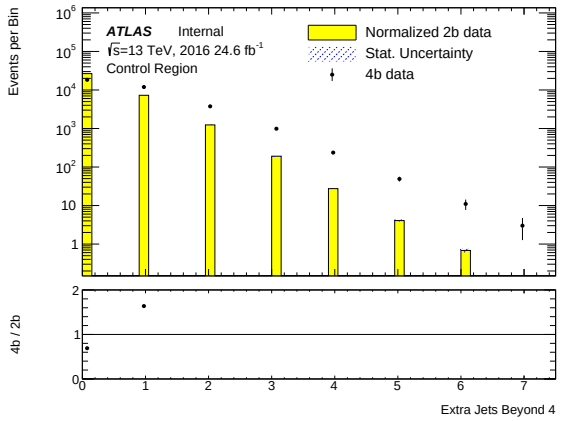
9.3 Reweighting Checks

In order to validate the performance of the BDT, the 2b and 4b distributions were compared for several key kinematic variables. Figures 9.3, 9.5, 9.7, 9.9, and 9.11 show the performance of the BDT reweighting in the control region while Figures 9.4, 9.6, 9.8, 9.10, and 9.12 show the performance in the validation region.

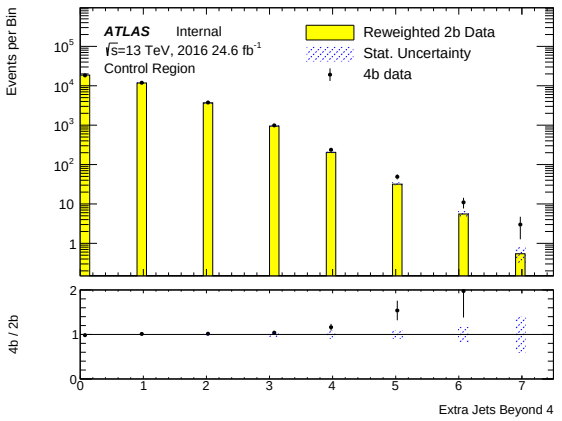
Good modeling is observed for all variables tested. The performance in the tails of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} distributions are difficult to gauge due to limited statistics, but seem generally consistent in the regions of interest. The highest $E_{\mathrm{T}}^{\mathrm{miss}}$ bin used for the analysis has a threshold of 200 GeV, while the highest m_{eff} bin has a threshold of 860 GeV.

9.4 Cross-Check with Simulation

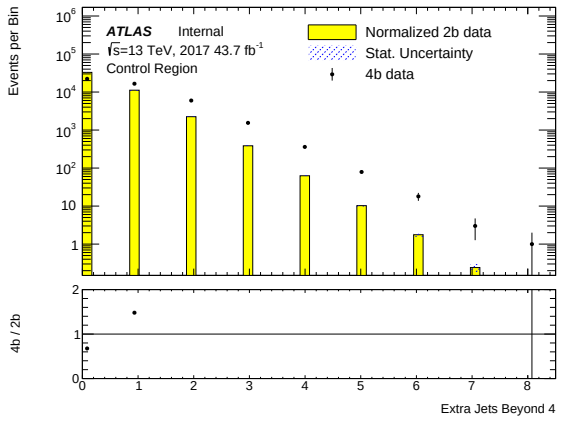
While the reweighting plots in Figures 9.3-9.12 show that the BDT reweighting procedure performs accurately on the dominant background of QCD multijet events, further checks were designed to ensure that the reweighting procedure is also able to learn the $t\bar{t}$ background. For these checks, the performance was evaluated using leptonic and semi-leptonic $t\bar{t}$ MC samples. The BDT was still trained on data using the 2b and 4b control regions as before, but the weights were now additionally predicted for 2b $t\bar{t}$ MC events. The reweighted 2b $t\bar{t}$ MC can then be compared to the 4b $t\bar{t}$ MC. Results are shown for each region in Figure 9.13. As QCD multijet dominates the background at low $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff}, $t\bar{t}$ has low statistics in these regimes and may not be able to be accurately modeled by the BDT. However, the contribution is small enough that the effects of any mismodeling would be negligible. Thus the plots only show the $t\bar{t}$ performance for high $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff}.



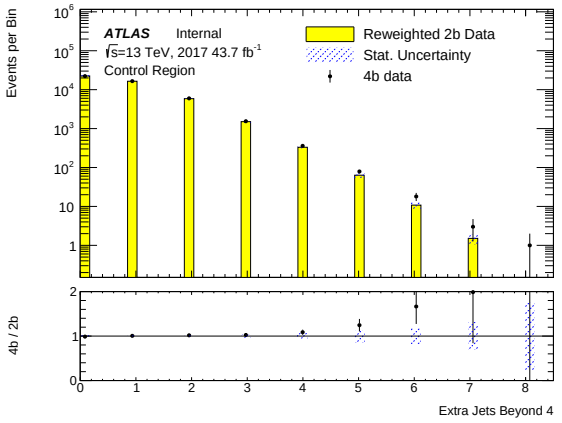
(a) 2016 unweighted



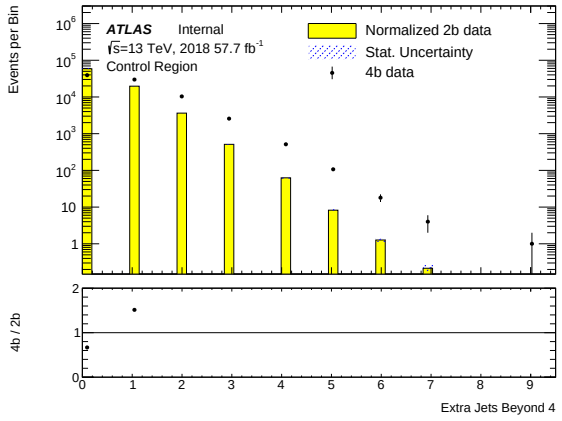
(b) 2016 reweighted



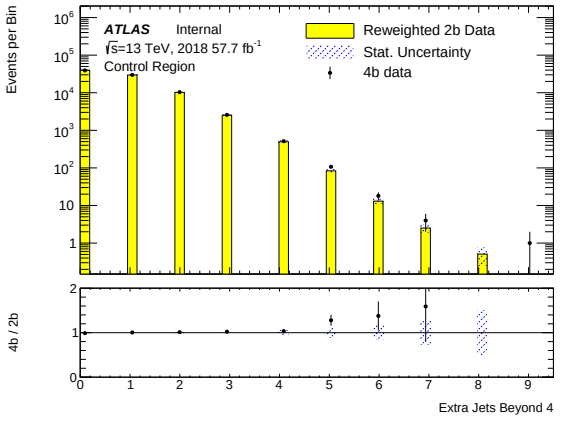
(c) 2017 unweighted



(d) 2017 reweighted

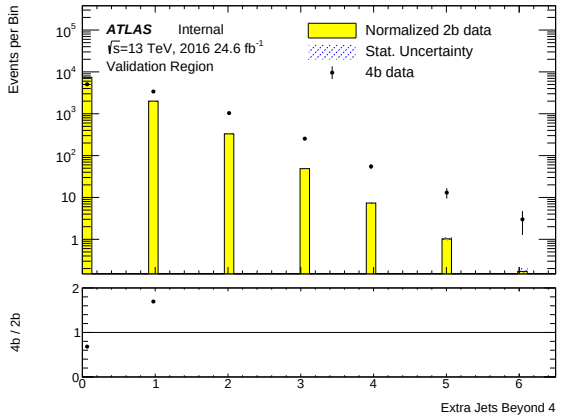


(e) 2018 unweighted

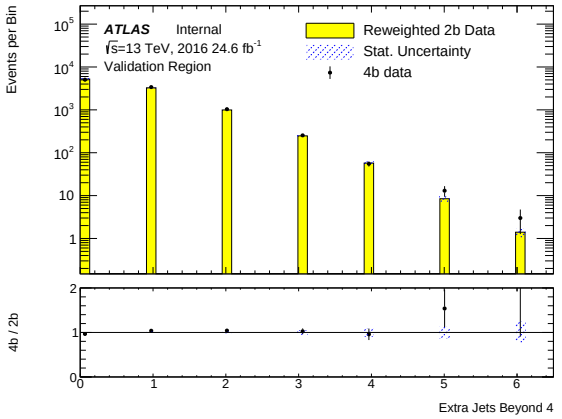


(f) 2018 reweighted

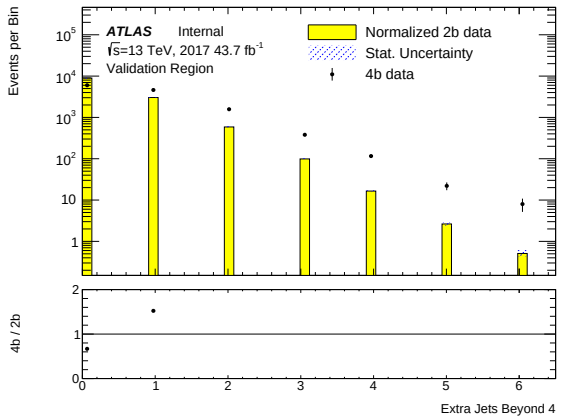
Figure 9.3: Comparisons of 2b and 4b number of jets distributions in the CR. The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics.



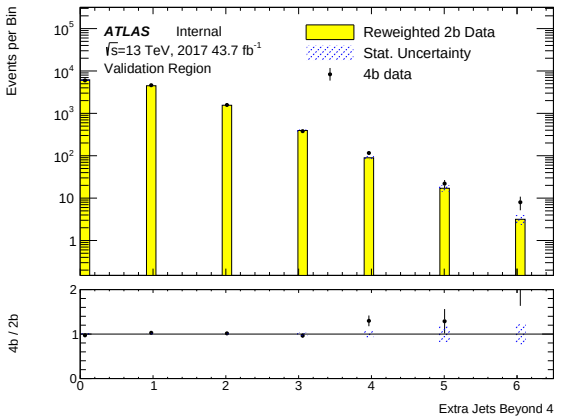
(a) 2016 unweighted



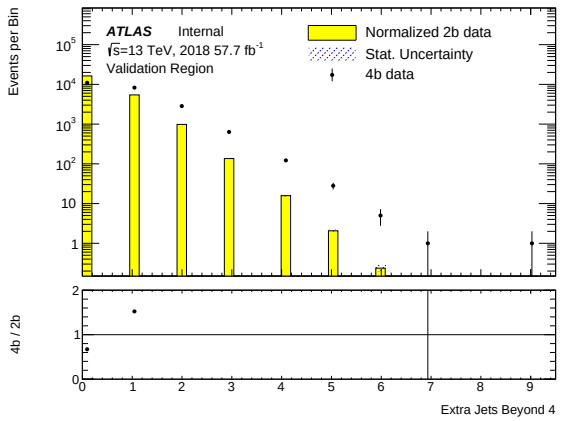
(b) 2016 reweighted



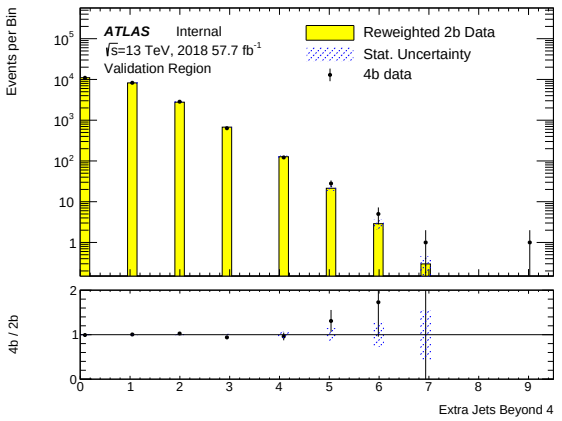
(c) 2017 unweighted



(d) 2017 reweighted

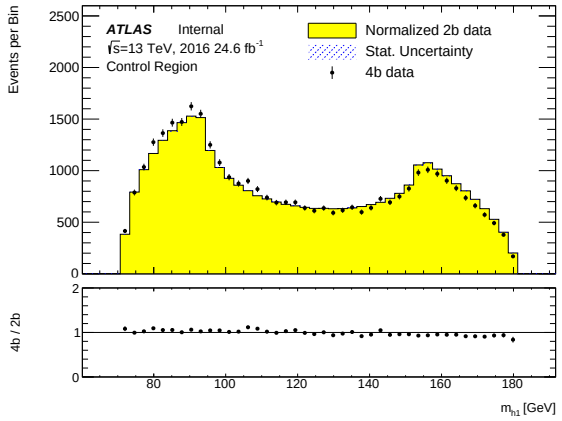


(e) 2018 unweighted

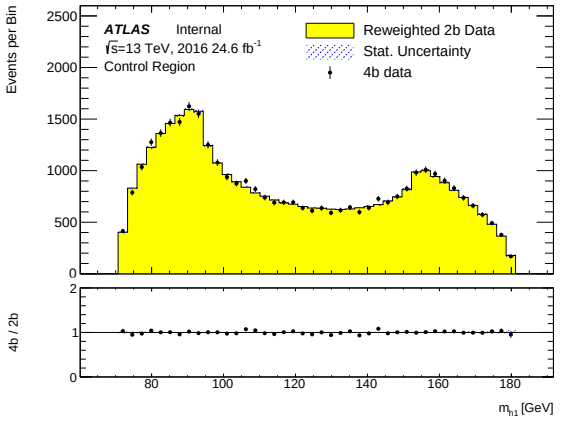


(f) 2018 reweighted

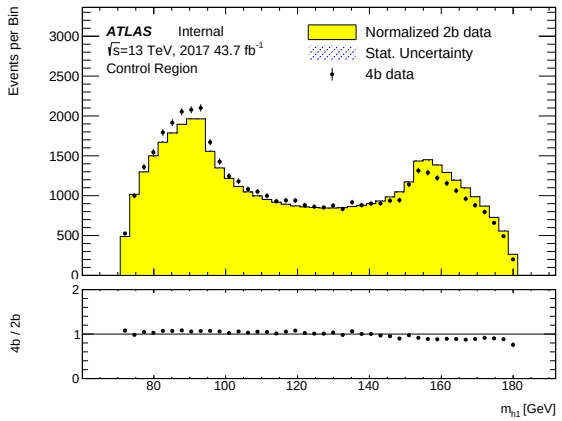
Figure 9.4: Comparisons of 2b and 4b number of jets distributions in the VR. The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics.



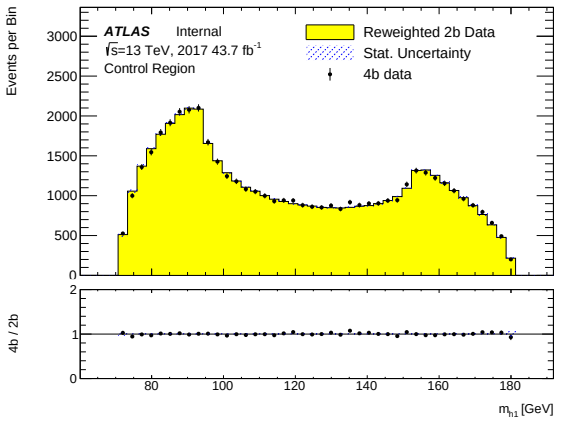
(a) 2016 unweighted



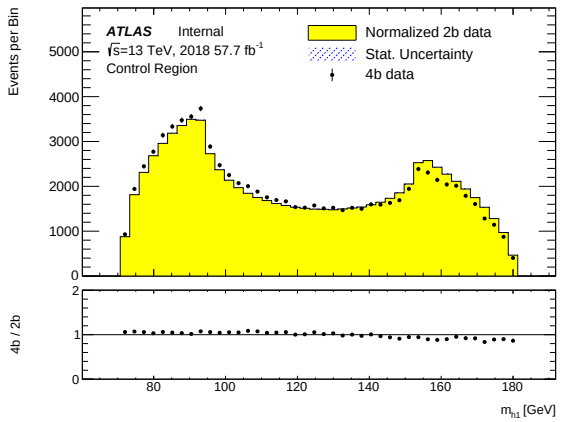
(b) 2016 reweighted



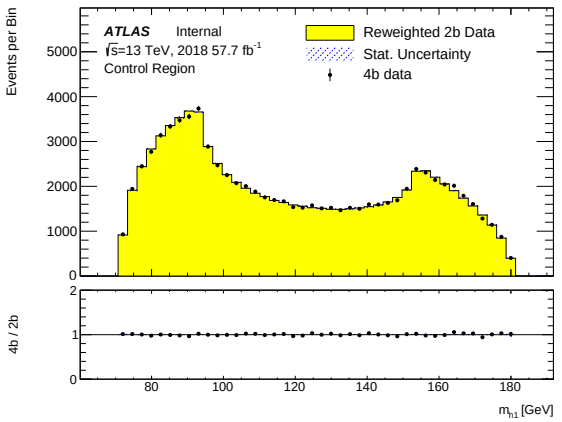
(c) 2017 unweighted



(d) 2017 reweighted

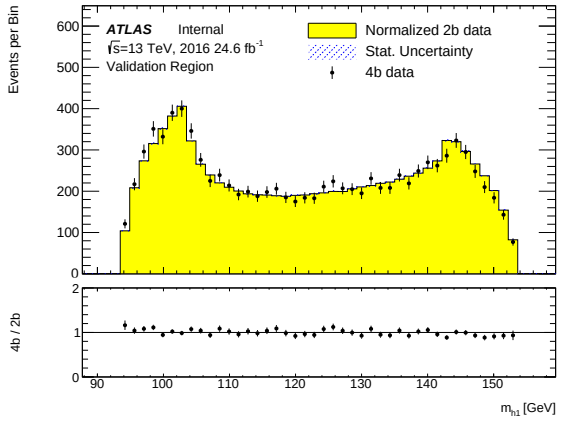


(e) 2018 unweighted

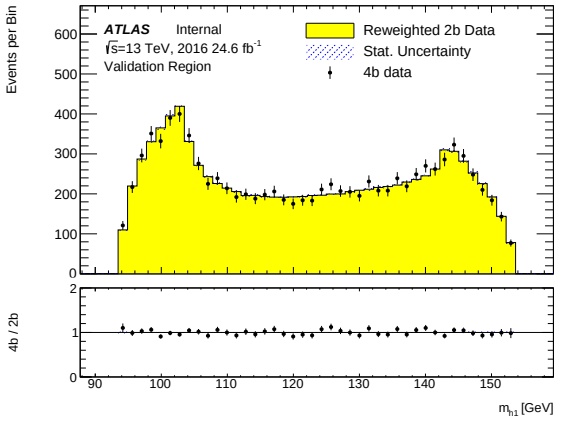


(f) 2018 reweighted

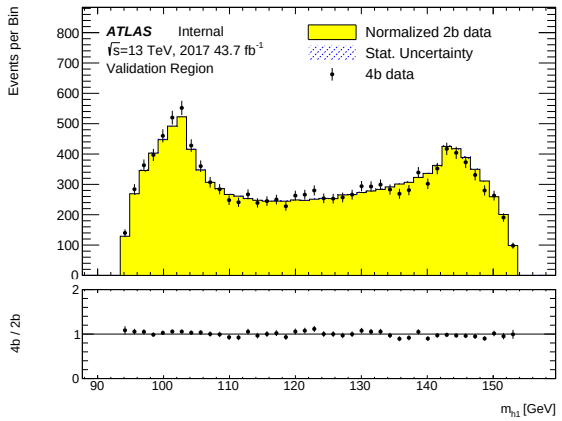
Figure 9.5: Comparisons of 2b and 4b leading Higgs candidate mass distributions in the CR. The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics.



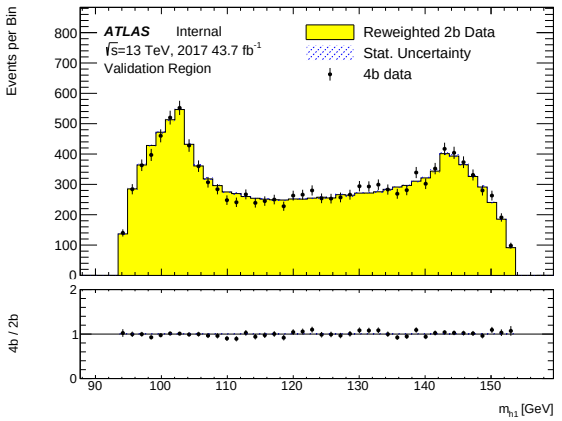
(a) 2016 unweighted



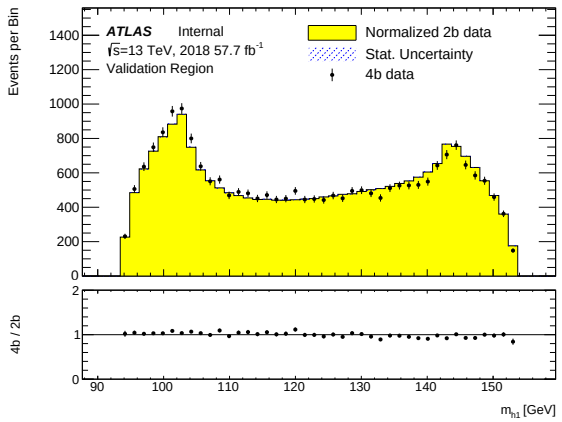
(b) 2016 reweighted



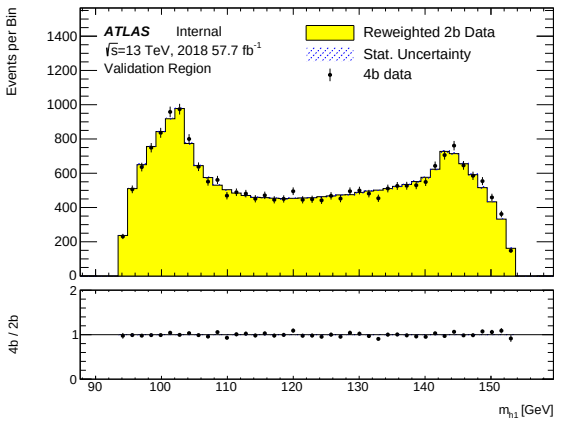
(c) 2017 unweighted



(d) 2017 reweighted

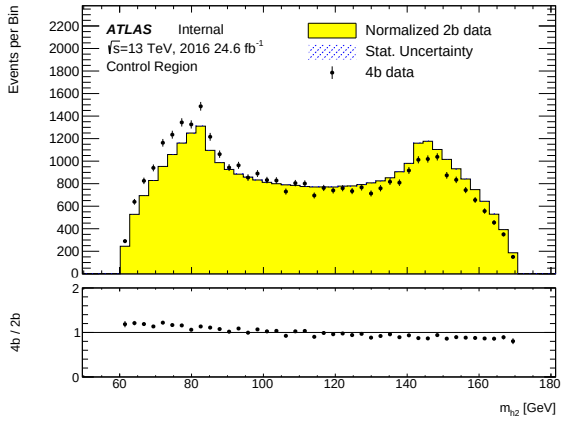


(e) 2018 unweighted

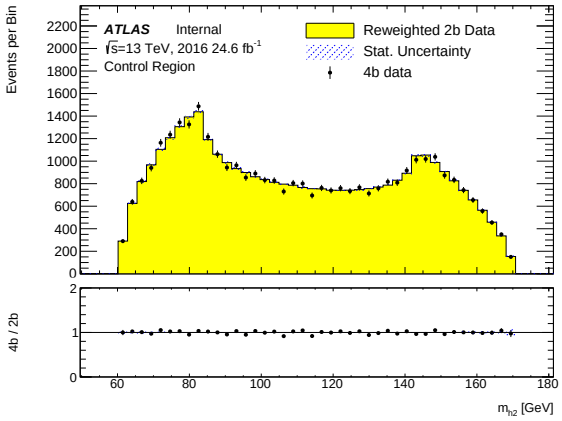


(f) 2018 reweighted

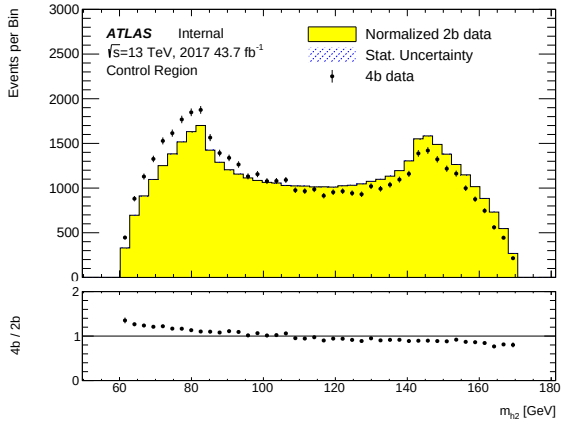
Figure 9.6: Comparisons of 2b and 4b leading Higgs candidate mass distributions in the VR. The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics.



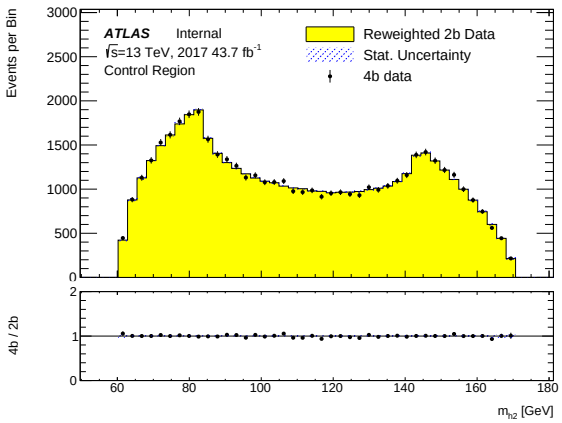
(a) 2016 unweighted



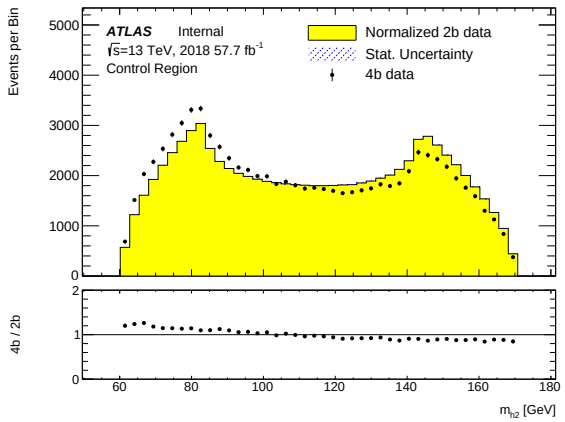
(b) 2016 reweighted



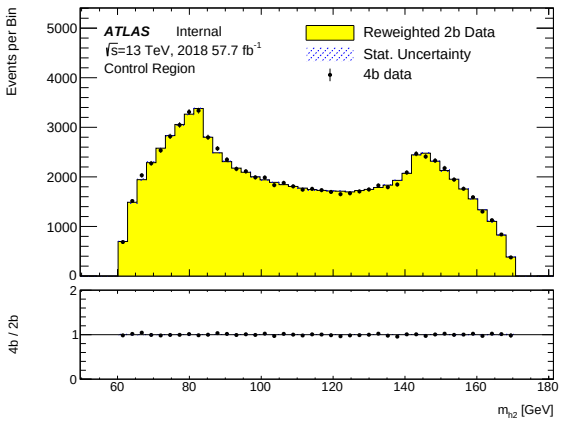
(c) 2017 unweighted



(d) 2017 reweighted

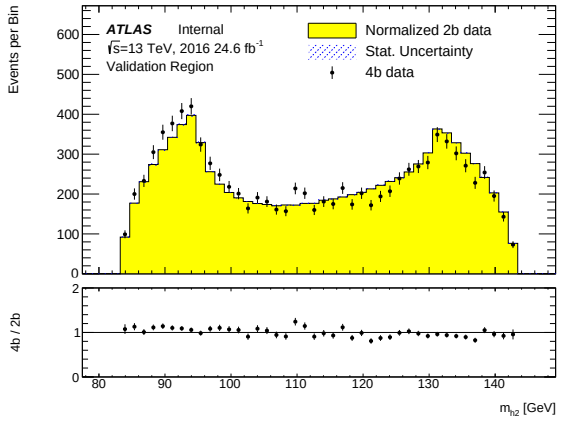


(e) 2018 unweighted

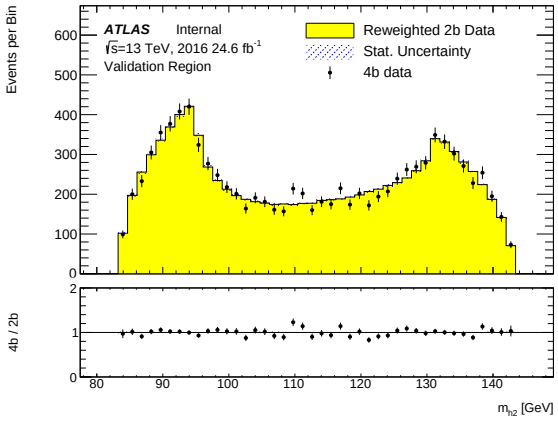


(f) 2018 reweighted

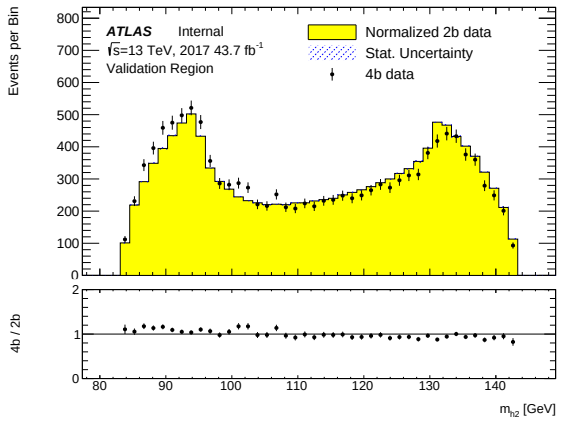
Figure 9.7: Comparisons of 2b and 4b sub-leading Higgs candidate mass distributions in the CR. The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics.



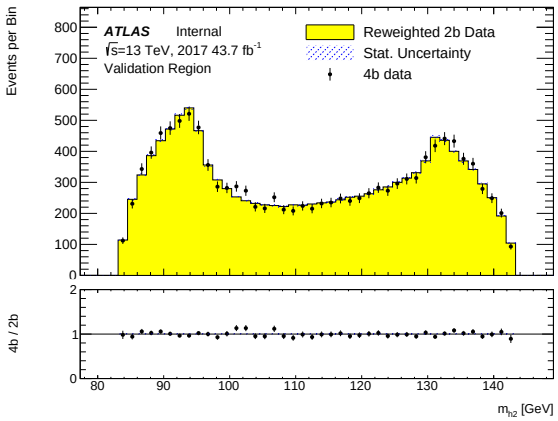
(a) 2016 unweighted



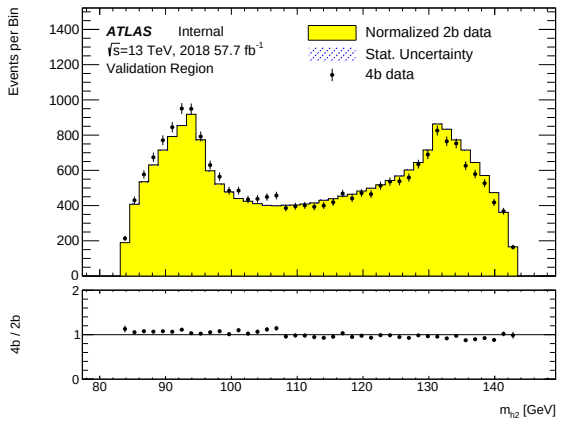
(b) 2016 reweighted



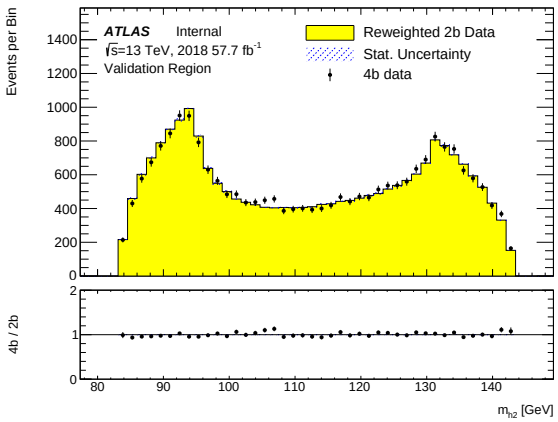
(c) 2017 unweighted



(d) 2017 reweighted

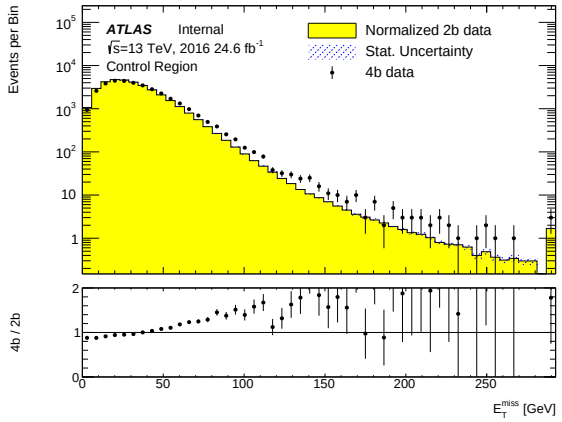


(e) 2018 unweighted

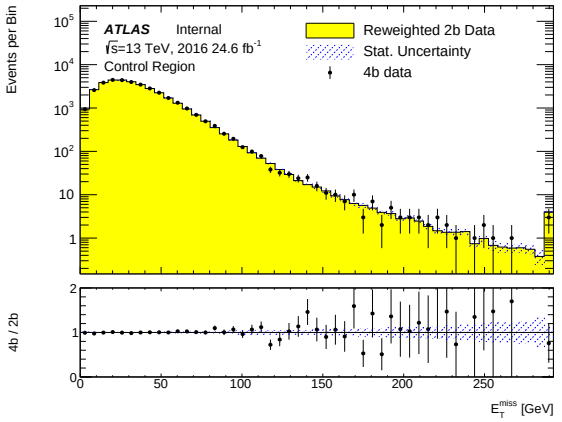


(f) 2018 reweighted

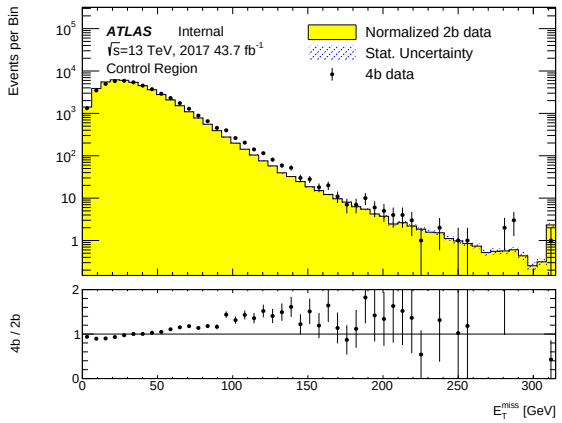
Figure 9.8: Comparisons of 2b and 4b sub-leading Higgs candidate mass distributions in the VR. The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics.



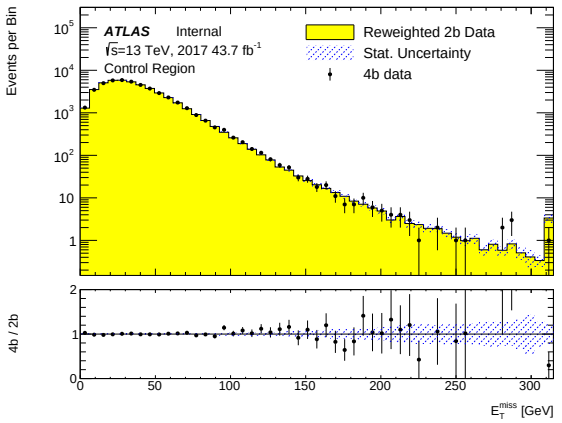
(a) 2016 unweighted



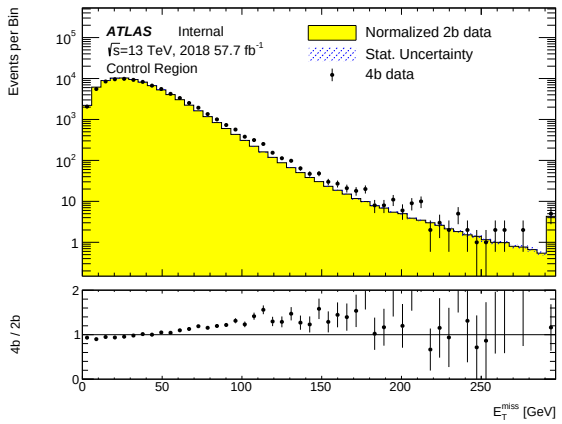
(b) 2016 reweighted



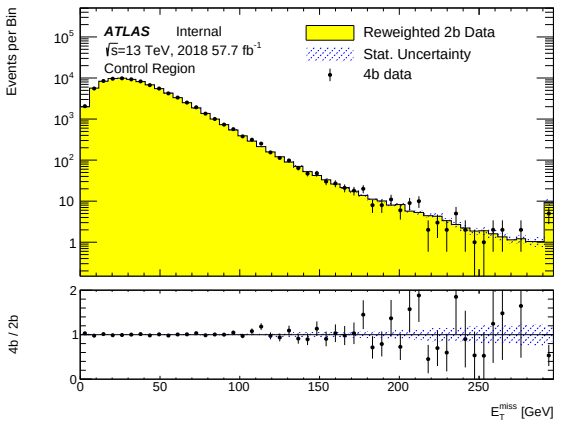
(c) 2017 unweighted



(d) 2017 reweighted

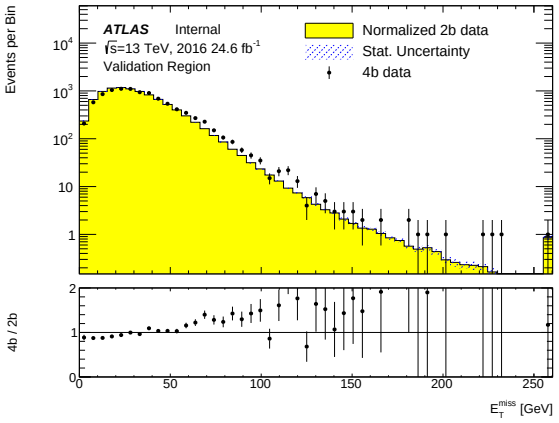


(e) 2018 unweighted

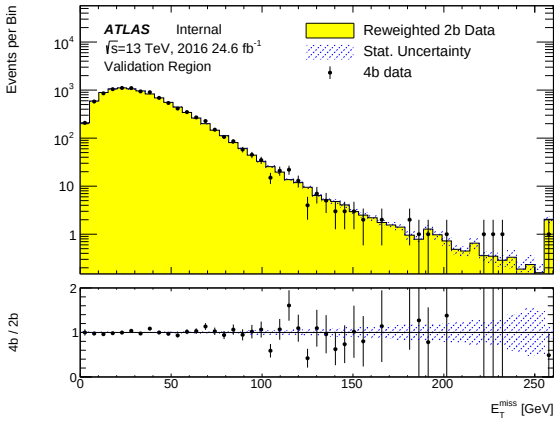


(f) 2018 reweighted

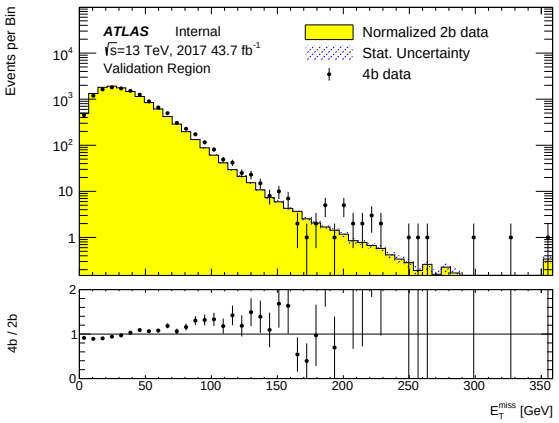
Figure 9.9: Comparisons of 2b and 4b $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the CR. The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics. The last bin shows overflow events.



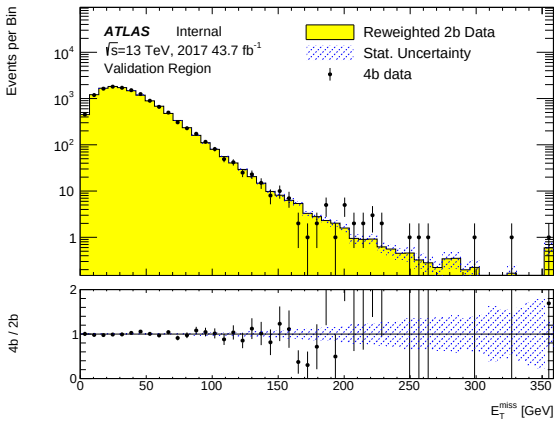
(a) 2016 unweighted



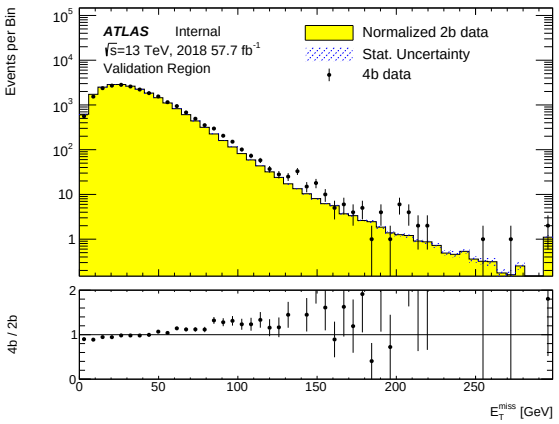
(b) 2016 reweighted



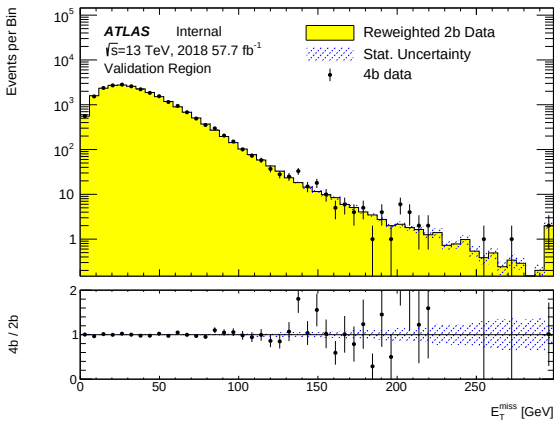
(c) 2017 unweighted



(d) 2017 reweighted

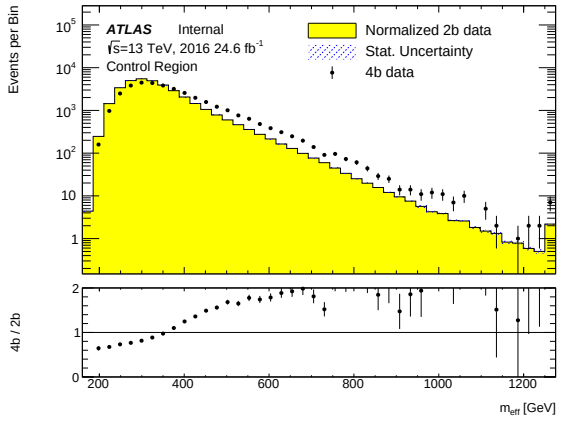


(e) 2018 unweighted

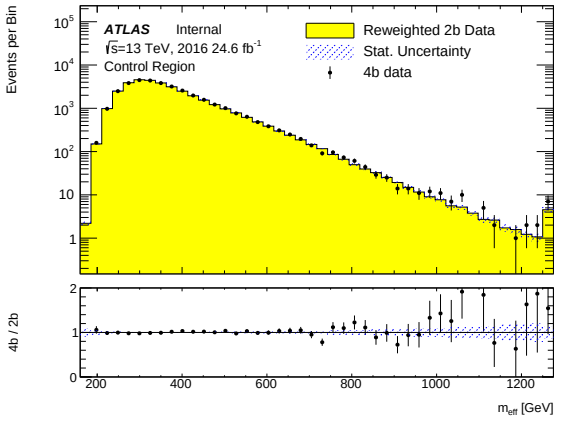


(f) 2018 reweighted

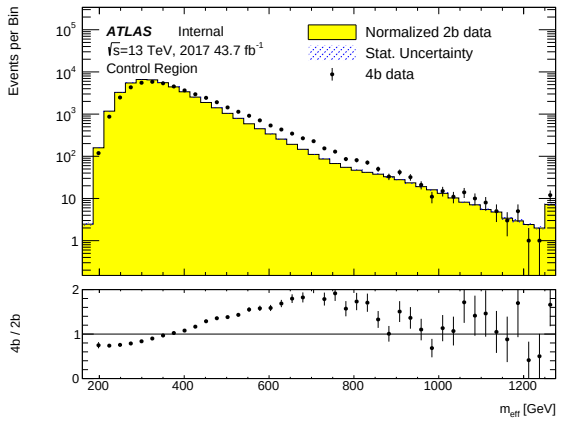
Figure 9.10: Comparisons of 2b and 4b $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the VR. The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics. The last bin shows overflow events.



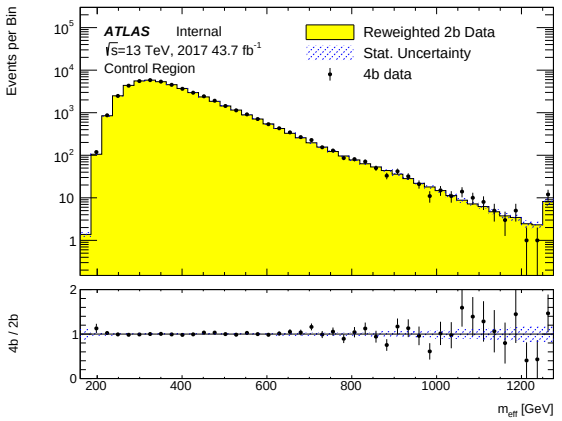
(a) 2016 unweighted



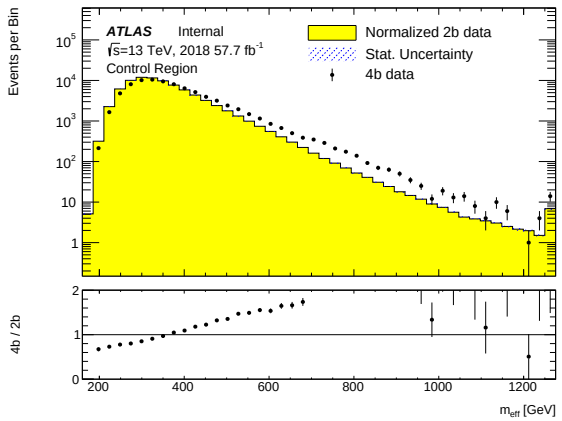
(b) 2016 reweighted



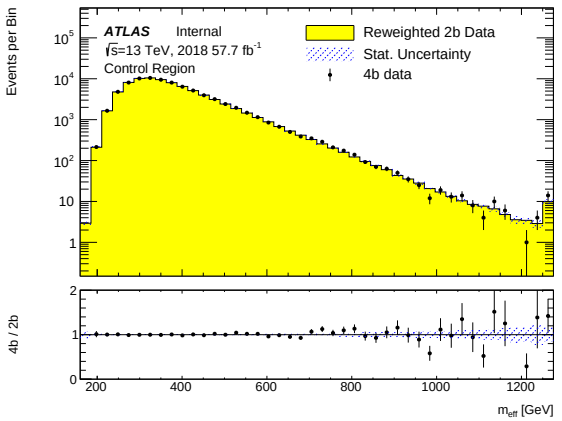
(c) 2017 unweighted



(d) 2017 reweighted

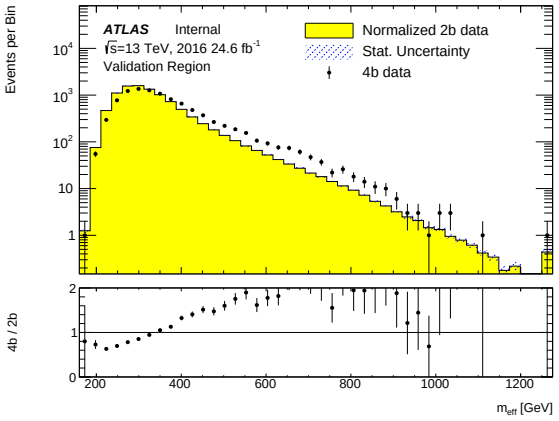


(e) 2018 unweighted

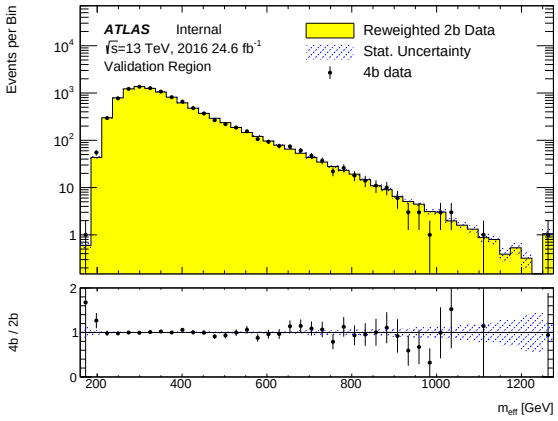


(f) 2018 reweighted

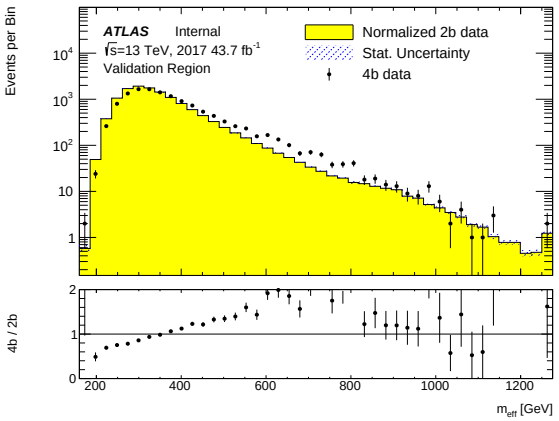
Figure 9.11: Comparisons of 2b and 4b m_{eff} distributions in the CR. The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics. The last bin shows overflow events.



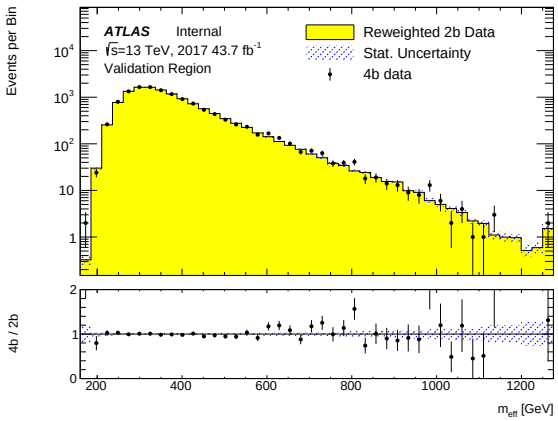
(a) 2016 unweighted



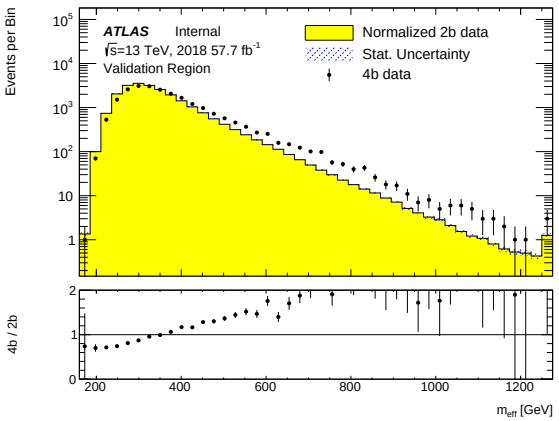
(b) 2016 reweighted



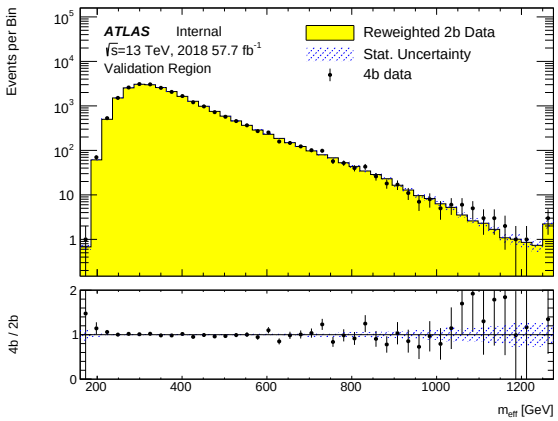
(c) 2017 unweighted



(d) 2017 reweighted

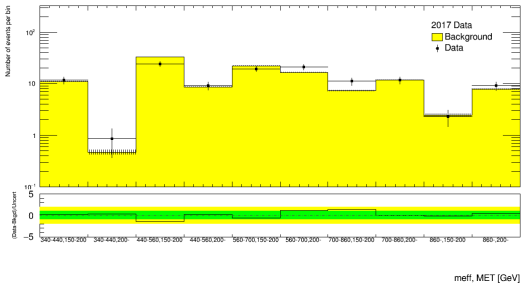


(e) 2018 unweighted

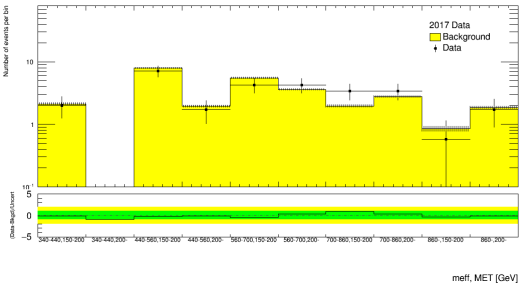


(f) 2018 reweighted

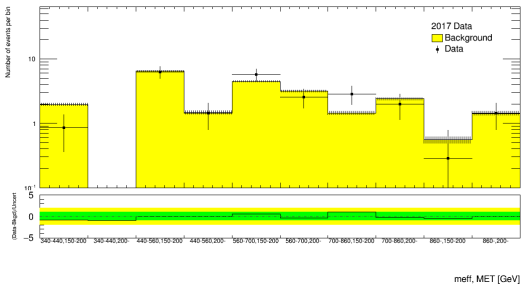
Figure 9.12: Comparisons of 2b and 4b m_{eff} distributions in the VR.The 2b distributions are normalized to the 4b integrals to show the shape differences. The blue shading shows the uncertainty due to the reweighting and 2b statistics. The last bin shows overflow events.



(a) Control Region



(b) Validation Region



(c) Signal Region

Figure 9.13: Reweighted 2b $t\bar{t}$ ("Background") vs. 4b $t\bar{t}$ ("Data") for 2017 $t\bar{t}$ MC samples. Events with $E_{\mathrm{T}}^{\mathrm{miss}} < 150$ GeV or $m_{\mathrm{eff}} < 340$ GeV are ignored as this regime is dominated by QCD multijet events, and any mismodeling due to the BDT reweighting performance on $t\bar{t}$ will be negligible.

Chapter 10

DISCOVERY AND EXCLUSION REGIONS

Limits are set on two types of scenarios. Model-dependent limits are set in terms of the cross section of the higgsino model under investigation. In addition, model-independent limits are set on the number of Beyond the Standard Model events.

10.1 Discovery Regions

The model-independent limits are set using dedicated discovery regions. Two such regions are constructed: one to target low-mass higgsinos and one to target intermediate-mass higgsinos. These are optimized for the $m_{\tilde{H}} = 150$ GeV and $m_{\tilde{H}} = 300$ GeV scenarios respectively. Each of these regions consists of a single bin with lower thresholds on $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff}.

10.1.1 Discovery Region Systematics

In addition to the statistical uncertainties (including the bootstrap uncertainty described in Chapter 11), two systematics are used. First, a non-closure uncertainty is needed to account for imperfections in the reweighting procedure in the control region. This is estimated as

$$\frac{N_{\mathrm{CR}}^{2b,\mathrm{rw}} - N_{\mathrm{CR}}^{4b}}{N_{\mathrm{CR}}^{2b,\mathrm{rw}}} \tag{10.1}$$

A second uncertainty is needed to account for differences in the control region (where the scale factor and weights are derived) and the signal region (where they are applied). This is set as

$$max\left(\frac{N_{\mathrm{VR}}^{2b,\mathrm{rw}} - N_{\mathrm{VR}}^{4b}}{N_{\mathrm{VR}}^{2b,\mathrm{rw}}}, \frac{N_{\mathrm{SR}}^{2b,\mathrm{CRrw}} - N_{\mathrm{SR}}^{2b,\mathrm{VRrw}}}{N_{\mathrm{SR}}^{2b,\mathrm{CRrw}}}\right) \tag{10.2}$$

Here VRrw refers to quantities calculated using an alternative set of weights derived in the validation region rather than the control region. The first of the two quantities above yields

the accuracy of the CR-derived weights at predicting the number of events in the validation region. The second quantity takes the difference between the estimates of the SR background using the CR-derived and VR-derived weights. The maximum is used as these both represent methods of accounting for the uncertainty due to extrapolating across the Higgs candidate mass plane.

The deviations for each uncertainty, given by the numerators in Equations 10.1 and 10.2, are calculated separately for the 2016, 2017, and 2018 data-taking periods. These are then combined in quadrature. This effectively weights the contribution from each year to correspond to the size of the dataset. These are then divided by the total counts over all years to convert to fractional uncertainties, which are then summed in quadrature.

10.1.2 Discovery Region Optimization

Various thresholds were tested for the $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} cuts used to define the discovery regions. No bootstraps were used in these tests. For $E_{\mathrm{T}}^{\mathrm{miss}}$, lower thresholds of 0, 20, 45, 70, 100, 150, and 200 GeV were considered, while for m_{eff}, lower thresholds of 160, 200, 260, 340, 440, 560, 700, and 860 GeV were considered. These correspond to the bin edges for the 2016 version of the analysis [26].

The significances for each region definition were calculated using the BinomialExpZ function of RooStats [92] and are shown in Figure 10.1. The optimal thresholds and the corresponding expected significances are shown in Table 10.1. The sizes of the uncertainties are shown in Table 10.2. If the optimization was done without systematics, the optimal thresholds for the 150 GeV region would be $E_{\mathrm{T}}^{\mathrm{miss}} > 0$ GeV and $m_{\mathrm{eff}} > 340$ GeV. While this has a large expected stat-only significance of 8.2σ, the expected significance with systematics would be only 3.2σ. Including the systematics in the optimization yields tighter cuts of $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and $m_{\mathrm{eff}} > 560$ GeV. While this reduces the stat-only expected significance to 7.0σ, it increases the full expected significance to 4.6σ. The optimal 300 GeV thresholds are much less dependent on systematic uncertainties, changing from $E_{\mathrm{T}}^{\mathrm{miss}} > 150$

Target $m_{\tilde{H}}$	$E_{\mathrm{T}}^{\mathrm{miss}}$ threshold	m_{eff} threshold	Expected σ (stat-only)	Expected σ (systematics)
130 GeV	20 GeV	560 GeV	9.5σ	6.2σ
150 GeV	20 GeV	560 GeV	7.0σ	4.6σ
200 GeV	150 GeV	340 GeV	7.5σ	6.6σ
250 GeV	150 GeV	340 GeV	8.7σ	7.7σ
300 GeV	150 GeV	340 GeV	8.4σ	7.4σ

Table 10.1: Thresholds and expected significances of the 150 GeV and 300 GeV discovery regions. Expected significances for other signal benchmarks are also shown, where each uses the region definition with the higher expected significance. 10 bootstraps are used for these estimates.

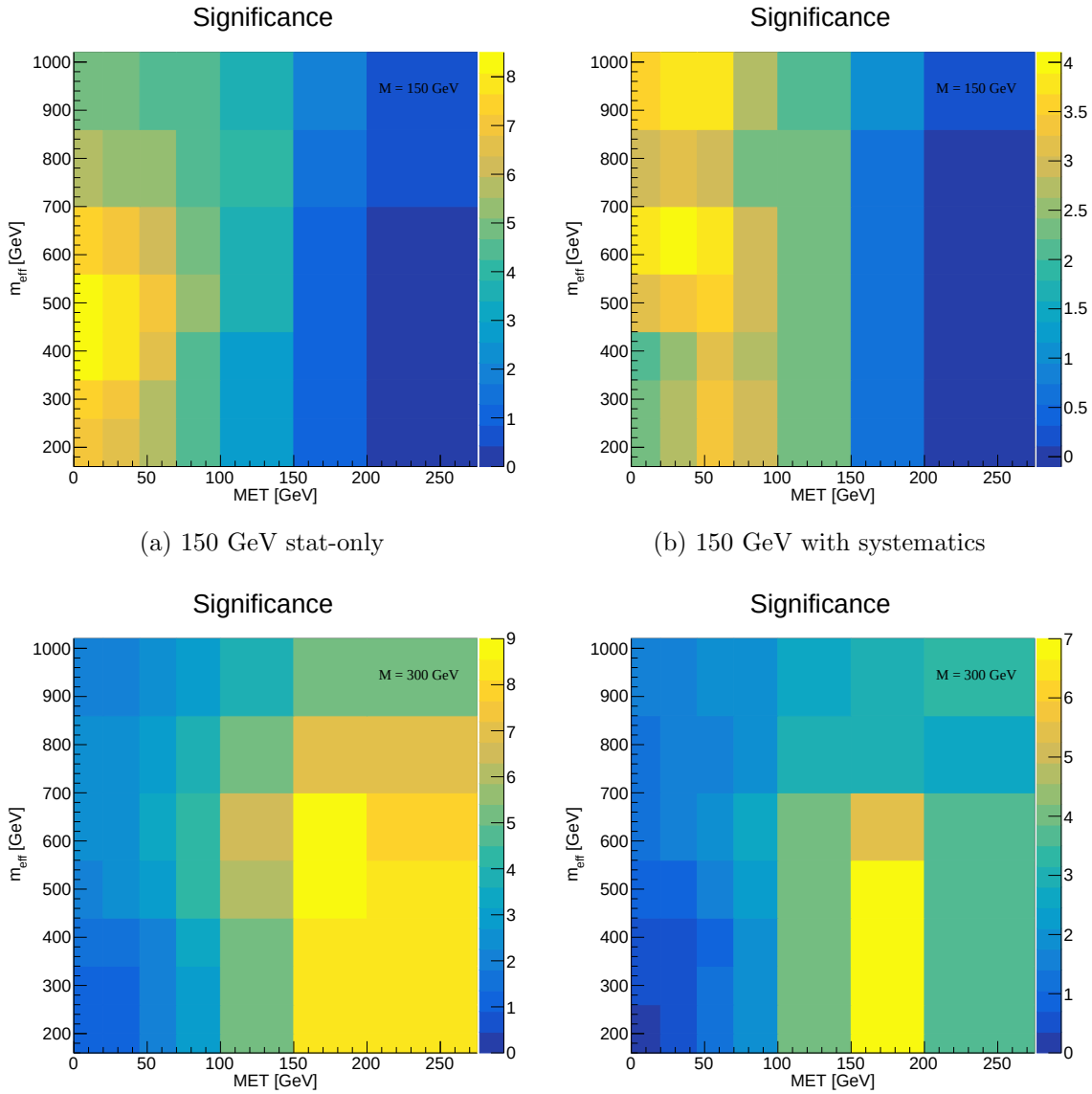
GeV and $m_{\mathrm{eff}} > 440$ GeV without systematics to $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV and $m_{\mathrm{eff}} > 340$ GeV with systematics. Higher mass signals are expected to be limited by statistics for two reasons. First, the cross section for higgsino production decreases as mass increases. Second, higher mass signals have higher typical $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} values, leading to lower backgrounds. The results for this region were consistent for all m_{eff} thresholds between 160 and 340 GeV. This is likely because the minimum m_{eff} is 160 GeV (for 4 jets) plus $E_{\mathrm{T}}^{\mathrm{miss}}$, yielding a minimum of 310 GeV. The 340 GeV value was selected to best match the optimal stat-only thresholds.

10.1.3 Discovery Region Evaluation

The final results with observations are obtained by setting limits with the CLs prescription (described in Chapter 13) using 20,000 pseudoexperiments. The fit is performed with a signal model of 1 event and no signal uncertainties, allowing limits to be set on the number of Beyond the Standard Model events.

10.2 Exclusion Region

If no significant excess is observed, model-dependent limits will be set using an exclusion region. These use the full $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} phase space. In order to maximize the sensitivity of



(a) 150 GeV stat-only

(b) 150 GeV with systematics

(c) 300 GeV stat-only

(d) 300 GeV with systematics

Figure 10.1: Expected discovery significances for 150 GeV and 300 GeV mass points for each $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} threshold under consideration. The left plots show the Poisson stat-only expected significances while the right plots show the expected significance with all systematics except bootstraps. The lower bin edges correspond to the thresholds being tested.

	150 GeV Bin Uncertainty	300 GeV Bin Uncertainty
Poisson	0.4%	1.9%
Bootstrap	1.1%	3.5%
$\frac{N_{\mathrm{CR}}^{2b,\mathrm{rw}}-N_{\mathrm{CR}}^{4b}}{N_{\mathrm{CR}}^{2b,\mathrm{rw}}}$	0.8%	2.1%
$\frac{N_{\mathrm{VR}}^{2b,\mathrm{rw}}-N_{\mathrm{VR}}^{4b}}{N_{\mathrm{VR}}^{2b,\mathrm{rw}}}$	2.1%	5.2%
$\frac{N_{\mathrm{SR}}^{2b,\mathrm{CRrw}}-N_{\mathrm{SR}}^{2b,\mathrm{VRrw}}}{N_{\mathrm{SR}}^{2b,\mathrm{CRrw}}}$	2.3%	2.5%
Total	2.7%	6.9%

Table 10.2: Background uncertainties for the 150 GeV and 300 GeV discovery regions. Values are expressed as a percent of the background yield. These values are before fitting, and summed in quadrature before being applied. Only the larger of $\frac{N_{\mathrm{VR}}^{2b,\mathrm{rw}}-N_{\mathrm{VR}}^{4b}}{N_{\mathrm{VR}}^{2b,\mathrm{rw}}}$ and $\frac{N_{\mathrm{SR}}^{2b,\mathrm{CRrw}}-N_{\mathrm{SR}}^{2b,\mathrm{VRrw}}}{N_{\mathrm{SR}}^{2b,\mathrm{CRrw}}}$ is used.

the analysis, the data were binned in both $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff}. The bins are defined using edges of

- $E_{\mathrm{T}}^{\mathrm{miss}}$: {0, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 13000} GeV
- m_{eff}: {160, 200, 260, 340, 440, 560, 700, 860, 13000} GeV

and are numbered as in Figure 10.2. These values are the finest possible binning subject to 3 conditions. First, the reweighting procedure must yield accurate CR and VR predictions in $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff}, as shown in Figures 9.9-9.12. This prevented the addition of extra bins at high $E_{\mathrm{T}}^{\mathrm{miss}}$ or m_{eff}. Second, the bins must be at least as wide as the estimated resolution. This would yield bin edges of

- $E_{\mathrm{T}}^{\mathrm{miss}}$: {0, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 13000} GeV
- m_{eff}: {160,195,230,270,315,360,410,465,520,580,645,710,780,860,13000} GeV

Finally, the bins are constrained by the available simulated signal statistics. If the bins are made too fine, bins with 0 signal events will appear between those with non-zero events. The

background statistics are sufficiently high that many of these bins will still have non-zero background events. This would artificially improve the expected limits as the background of the signal-free bins would effectively be thrown out in the fit. Requiring minimal such gaps widens the bins to the final values.

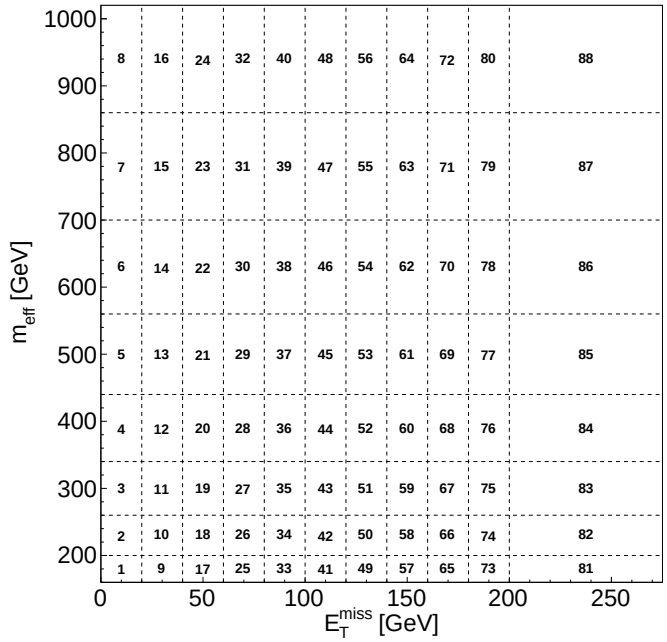


Figure 10.2: Bin-mapping for 2D unrolled distributions. The highest $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} bins have an upper threshold of 13 TeV.

Chapter 11

SYSTEMATIC UNCERTAINTIES

This analysis has five types of systematic uncertainties. First, there are uncertainties on the modeling of our triggers in Monte Carlo simulation, as discussed in Chapter 8. Second, there are uncertainties on the data-driven background estimate, relating to imperfections in the estimation method. The next category of uncertainties is those due to non-trigger detector mismodeling in our Monte Carlo simulations. As our background estimate is purely data-driven, these only affect our signal Monte Carlo samples. We are also affected by theoretical uncertainties in the signal Monte Carlo generation. Finally, an uncertainty of 1.7% is assigned to the luminosity.

11.1 Background Modeling

The background model has several sources of statistical and systematic uncertainties. First, limited statistics in the control region affect the derivation of the BDT weights, giving the final distributions uncertainties higher than simple Poisson statistics. Second, the reweighting procedure is imperfect, with the reweighted 2b control region not matching the 4b control region exactly. Finally, the ABCD procedure has imperfections corresponding to correlations between the region definitions and the number of b-tagged jets. This last source of uncertainty is split into two parts: one to account for the validity of applying the CR-derived normalization factor in the SR (transfer normalization uncertainty), and one to account for the validity of applying the CR-derived weights in the SR (transfer shape uncertainty). Each of these systematics is derived separately for the 2016, 2017, and 2018 data-taking periods.

The statistical uncertainty of the BDT training is evaluated using a bootstrapping procedure. To account for how the input statistics affect the training of the BDT, random Poisson weights are applied to each input event in both the 2b and 4b samples. The BDT is then

re-derived using these Poisson-weighted events. The weight-prediction step is performed using standard (unweighted) 2b events. This procedure is repeated for 100 different sets of Poisson weights, creating 101 BDT variations (including the nominal BDT using unweighted events). Whenever we draw a distribution, we set the value in each bin to the median of the values predicted by each of these 101 BDTs. We then set half of the bin-by-bin difference between the 84th and 16th percentiles of these 101 estimates as a bootstrap uncertainty. This gives us an estimate of how much the weights vary under statistical fluctuations of the inputs. The total statistical uncertainty on 2b samples is set to the quadrature sum of the bootstrap uncertainty and the Poisson uncertainty.

The imperfections in the BDT reweighting are referred to as the non-closure uncertainty. This is given by the differences between the 4b and reweighted 2b samples shown in Figure 11.1. For each bin, the fractional uncertainty is set as an uncorrelated shape systematic. The difference between the 4b and reweighted 2b samples can come from two sources: non-closure of the reweighting procedure and statistical uncertainties. In order to avoid double-counting the latter, no systematic is set for bins where the difference is less than the combined statistical (1σ) uncertainty as determined using the bootstrapping procedure.

If there are correlations between the region definition (from the location in the Higgs candidate mass plane) and the number of b-jets, the scale factors and kinematic corrections will depend on the region. A transfer shape systematic is defined to account for these differences. There are two possible ways to estimate this. The accuracy of the reweighting can be measured in the VR by taking the bin-by-bin difference between the 4b sample and the reweighted 2b sample. In contrast to its behavior in the discovery regions (see Chapter 10), this non-closure (Figure 11.2) is of similar size to that in the CR and consistent with the expected pulls for purely statistical variations. Therefore, an alternative method is used to estimate the systematic uncertainty.

The second method accounts for the stability of the weights with respect to the region in which they are derived. As in Chapter 10, the weights are re-derived using the validation

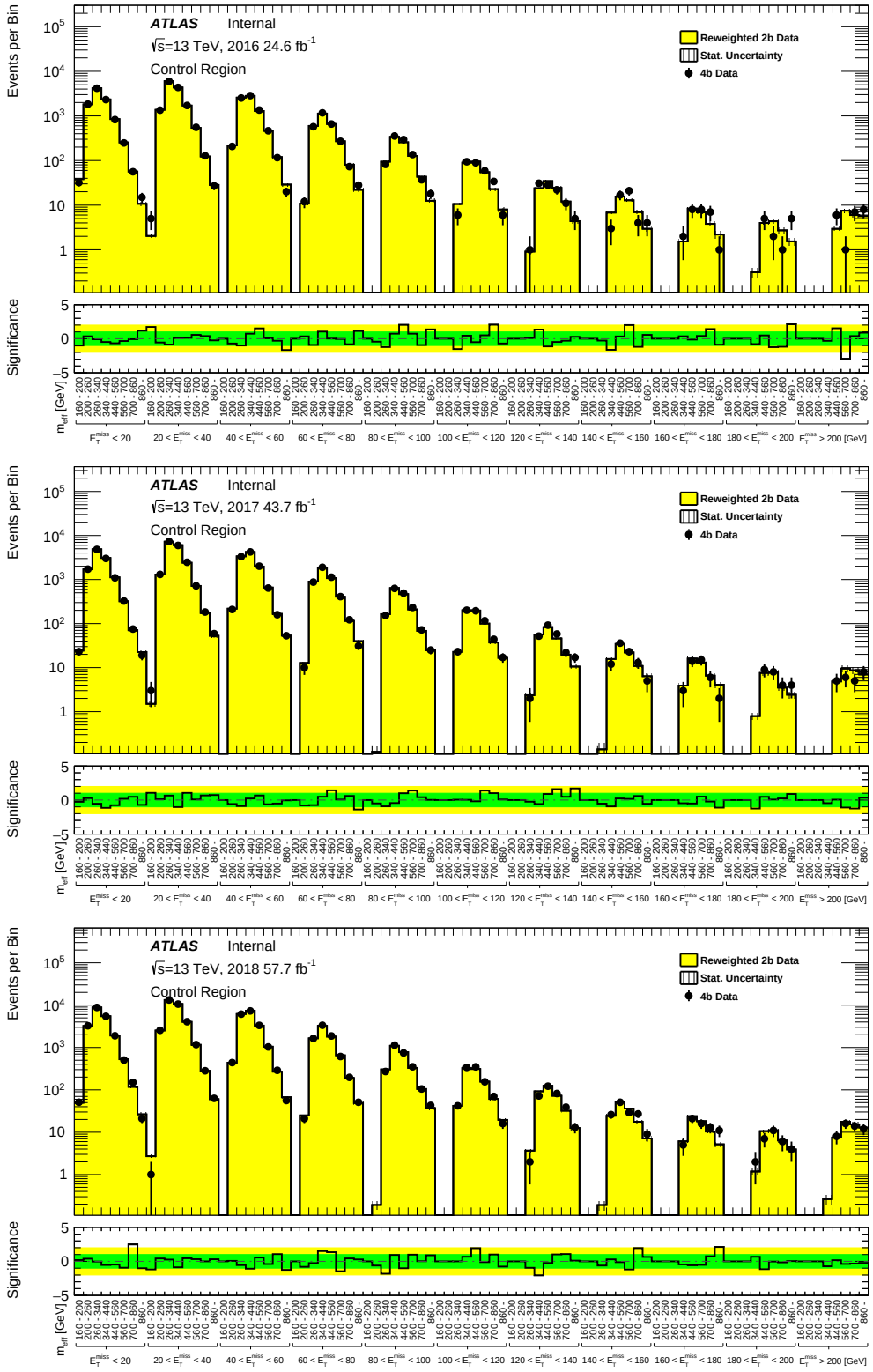


Figure 11.1: Shape differences between the 4b and reweighted 2b distributions in the 2016, 2017, and 2018 CRs. Due to the definition of the weights, the 2b and 4b distributions have the same normalization.

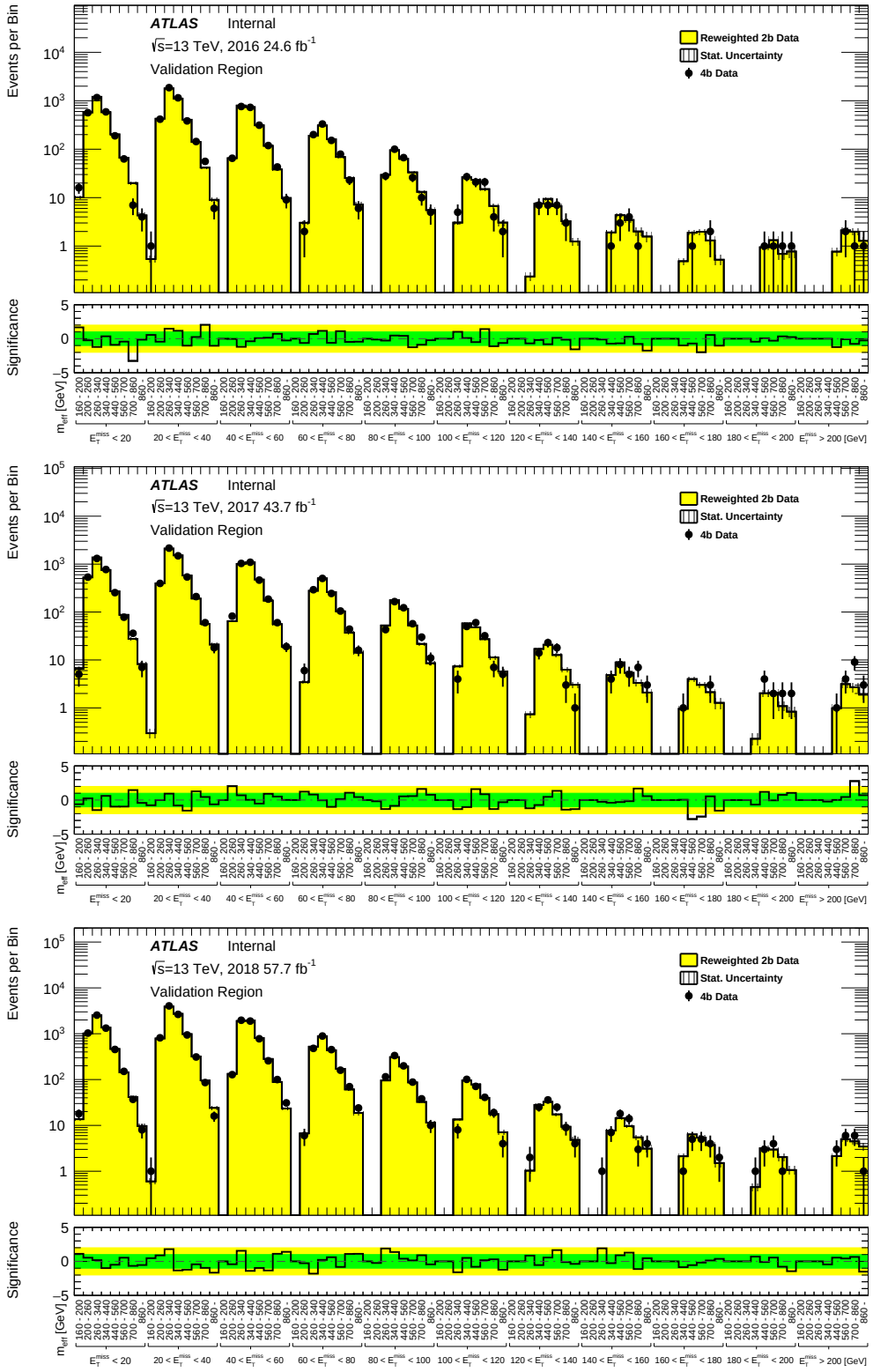


Figure 11.2: Shape differences between the 4b and reweighted 2b distributions from the CR-derived weights in the 2016, 2017, and 2018 VRs. The 2b distributions are normalized to the 4b distributions.

region instead of the control region. The full bootstrapping procedure is applied. The VR-derived estimate is then normalized to the CR-derived estimate as the normalization uncertainty is accounted for separately. The bin-by-bin difference of the two estimates in counts is then set as an uncorrelated shape systematic (Figure 11.3). The procedure is similar to the non-closure estimate; in order to avoid double-counting uncertainties, no systematic is set for bins where the difference between the predictions is within the statistical uncertainties.

Finally, we need the uncertainty on the normalization of the background estimate. The nominal value could be incorrect for the same reasons as the transfer shape uncertainty, so a similar procedure is applied. First, the accuracy of the normalization can be measured in the VR by taking the percent difference of $N_{\mathrm{VR}}^{2b,\mathrm{rw}} = \bar{w}_{\mathrm{VR}} * \mu_{\mathrm{CR}} * N_{\mathrm{VR}}^{2b}$ and N_{VR}^{4b}, where $\bar{w}$ represents the average weight in a given region. Second, the stability of the normalization with respect to weight derivation region can be estimated. To do this, the scale factor and weights are re-derived using the validation region instead of the control region. The uncertainty is then given by the percent difference of the two predictions in the SR, $N_{\mathrm{SR}}^{2b,\mathrm{CRrw}} = \bar{w}_{\mathrm{SR}}^{\mathrm{CRrw}} * \mu_{\mathrm{CR}} * N_{\mathrm{SR}}^{2b}$ and $N_{\mathrm{SR}}^{2b,\mathrm{VRrw}} = \bar{w}_{\mathrm{SR}}^{\mathrm{VRrw}} * \mu_{\mathrm{VR}} * N_{\mathrm{SR}}^{2b}$. The maximum of these two estimates is set as a normalization systematic uncertainty.

The non-closure and transfer shape systematics were also tested with systematics correlated across $E_{\mathrm{T}}^{\mathrm{miss}}$ or m_{eff} bins. It was found that the uncorrelated scheme provided the most conservative limits. In the absence of a rigorous argument for how the systematics should be correlated, the conservative approach was used as it was deemed safest.

The relative contributions of each uncertainty are shown in Figure 11.4.

11.2 Detector Modeling

Detector modeling uncertainties reflect differences between the simulated and actual behavior of the ATLAS detector. As the background estimate is entirely data-driven, these only affect the signal samples. These are further divided into two types. Up and down vari-

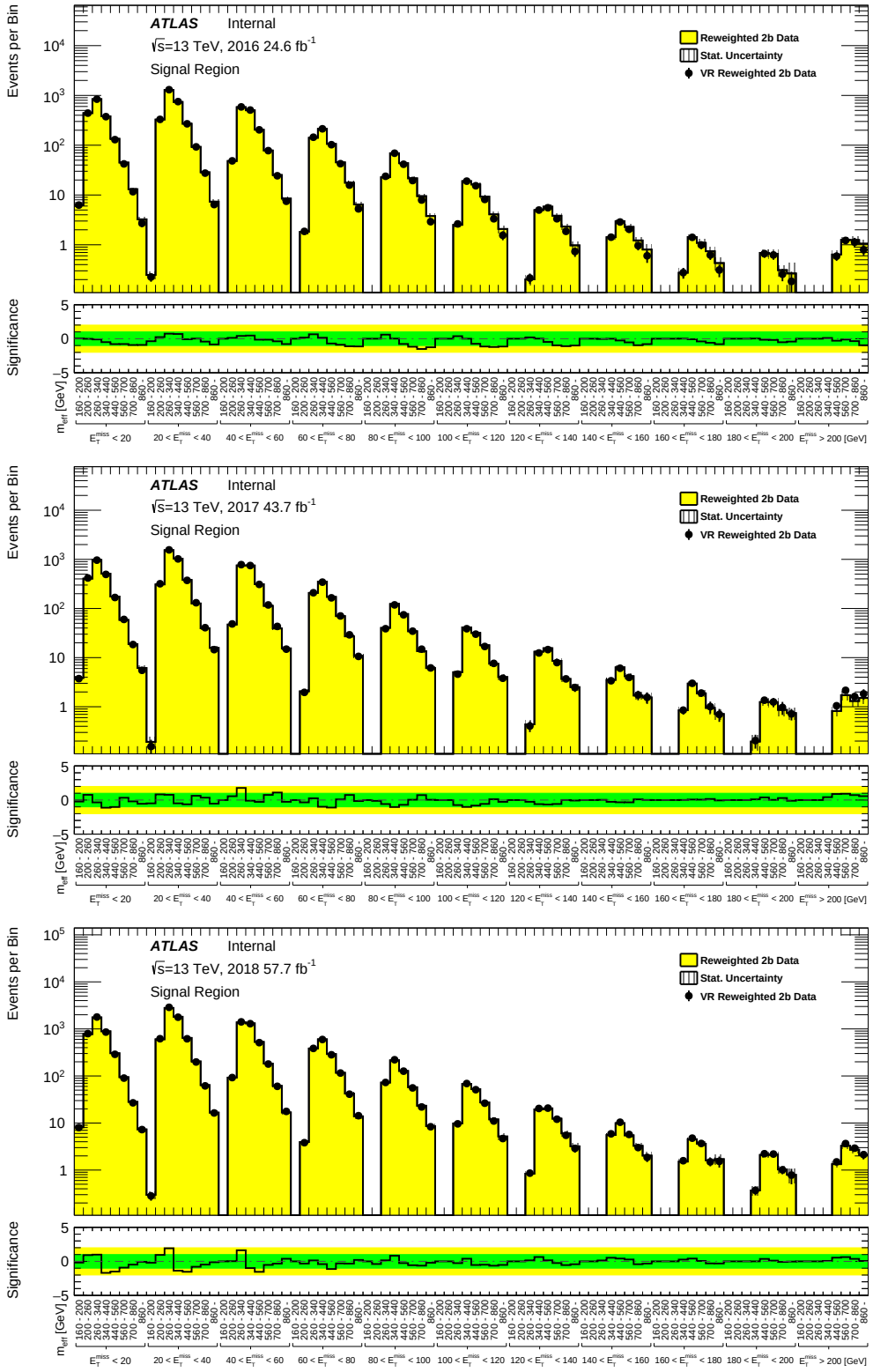


Figure 11.3: Shape differences between the reweighted 2b distributions from the CR-derived and VR-derived weights in the 2016, 2017, and 2018 SRs. The VR-derived distributions are normalized to the CR-derived distributions.

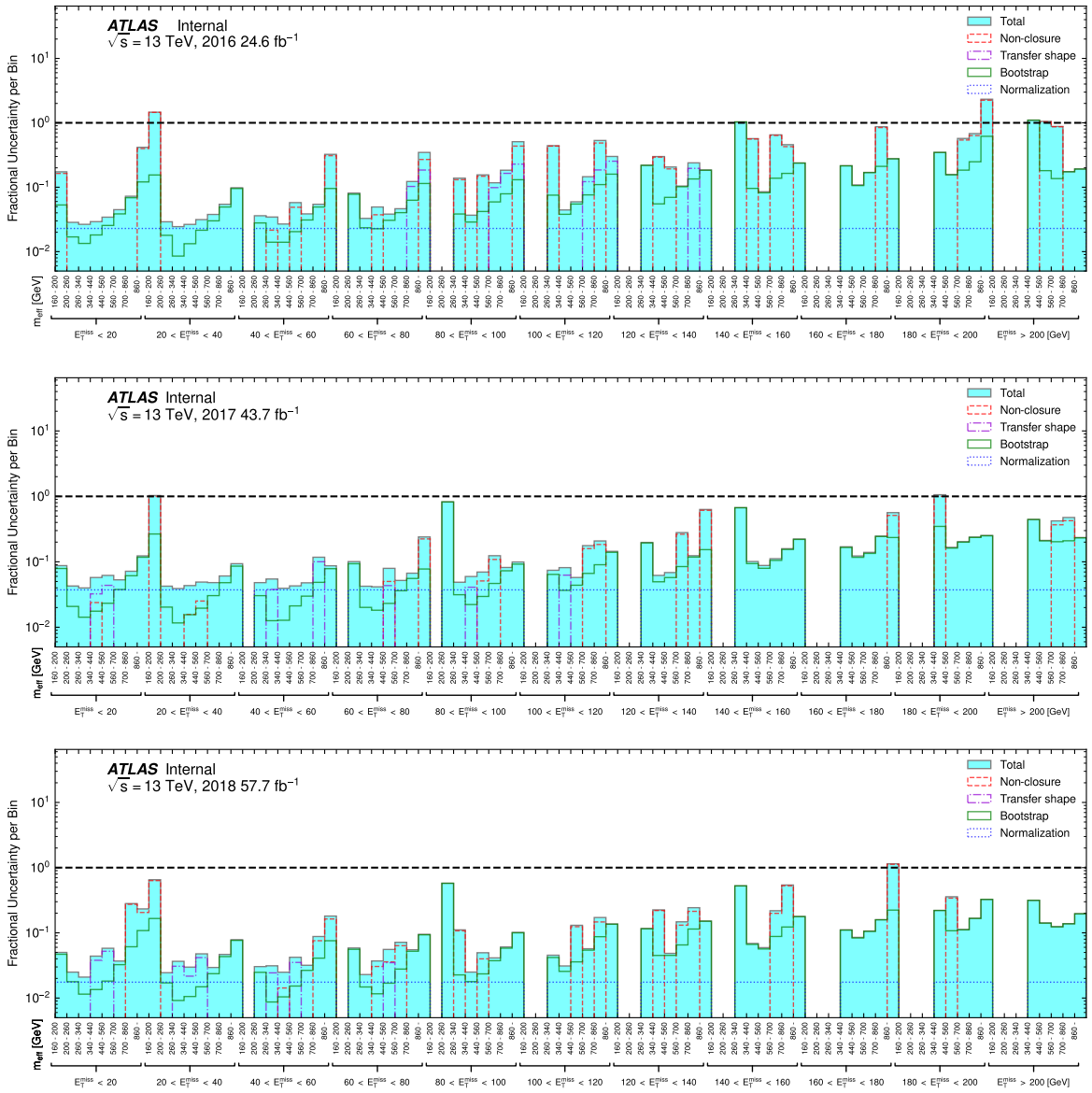


Figure 11.4: Background uncertainties as a fraction of background yields for the 2016, 2017, and 2018 SRs. Pre-fit values are shown. The bootstrap component also includes the Poisson component of the statistical uncertainty on the 2b sample. The total uncertainty shown is the quadrature sum of the individual uncertainties. All uncertainties are treated as symmetric. Values are not shown for bins with 0 background events, though there is a small uncertainty from the 2b Poisson uncertainty.

ations for each of these are applied as shape systematics, correlated between years. These shapes are varied coherently rather than on a bin-by-bin basis.

The first type is uncertainties on scale factors. These scale factors are event-level weights used to correct MC, and have associated uncertainties. There are 19 nuisance parameters (NPs) related to b-tagging: 9 for the b-tagging efficiency, 4 for the c-quark mis-tagging, 4 for the light quark mis-tagging, and 2 for extrapolating the b-jet and c-jet weights to high p_{T}. These are combined with the b-jet trigger modeling uncertainties as discussed in Chapter 8. The remaining scale factor uncertainties are from the jet vertex tagger (JVT), pile-up reweighting (PRW), and jet/H_{T} trigger scale factors. The JVT is an ATLAS software tool for rejecting pile-up jets [88].

The second set of uncertainties are on jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. These are the jet energy resolution (JER), jet energy scale (JES), jet mass scale (JMS), and $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term uncertainties.

The JES calibration rescales the four-momenta of jets to match truth-level Monte Carlo jets. This procedure corrects for effects such as pile-up and energy leakage. Many components of this calibration have uncertainties, leading to 125 independent nuisance parameters (NPs). The number of NPs is reduced to 30 using the "CategoryReduction" scheme. This scheme groups the p_{T}-dependent *in-situ* uncertainties into those arising from detector effects, statistical effects, modeling effects, or a combination thereof. Each group is decomposed into orthogonal terms, with larger terms set as NPs while smaller terms are combined [93].

The JER is the uncertainty on individual jet energies. This can be written as

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{N}{p_{\mathrm{T}}} \oplus \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus C \tag{11.1}$$

where N is a noise term due to electronics and pile-up, S is a stochastic term due to statistical variations, and C is a constant term due to various effects [93]. The JER is measured by calculating the asymmetry between jets in well-understood dijet samples. Systematic uncertainties include non-closure of the calculations, propagation of JES uncertainties, and

variations of MC generators and selections [93].

Due the to the limited statistics in some of our bins, the JER systematics can be rather large. This is mitigated by using a smoothing algorithm with a Gaussian kernel based on the procedure described in [94]. The JER prescription is used to generate distributions for the nominal, up, and down variations. Any bins with negative values are then set to 0. For each bin i, we define

$$c_{\mathrm{up},i} = N_{\mathrm{up},i}/N_{\mathrm{nom},i} \tag{11.2}$$

and similar for $c_{\mathrm{down},i}$. We then calculate new bin values using

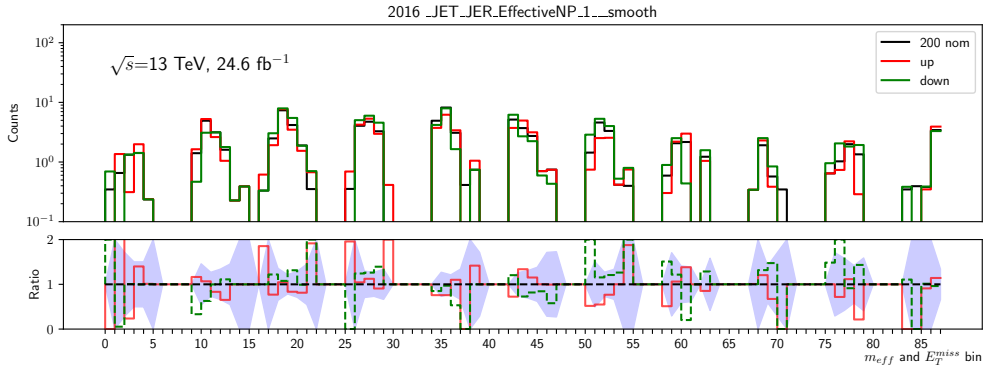
$$\tilde{N}_{\mathrm{up},j} = N_{\mathrm{nom},j}\frac{\sum_i w_{ij} * c_i}{\sum_i w_{ij}} \tag{11.3}$$

where w_{ij} is given by

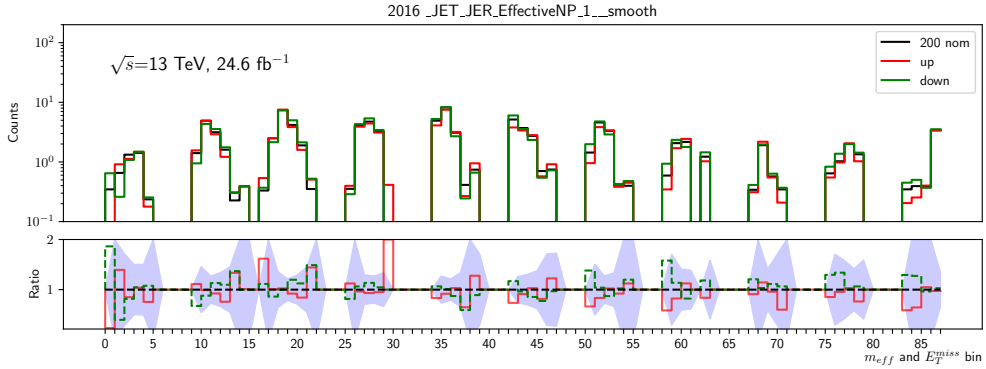
$$w_{ij} = N_{\mathrm{nom,noweights},i} Exp(-(\Delta_{ij} E_{\mathrm{T}}^{\mathrm{miss}})^2/2\sigma_{\mathrm{met}}^2 - (\Delta_{ij} m_{\mathrm{eff}})^2/2\sigma_{\mathrm{meff}}^2) \tag{11.4}$$

The factor $N_{\mathrm{nom,noweights},i}$ is the unweighted version of $N_{\mathrm{nom},i}$. This gives higher weight to bins that have more statistics (and therefore lower uncertainties). $\Delta_{ij} E_{\mathrm{T}}^{\mathrm{miss}}$ ($\Delta_{ij} m_{\mathrm{eff}}$) is the difference between the mean $E_{\mathrm{T}}^{\mathrm{miss}}$ (m_{eff}) values for bin i and bin j, where the mean for the highest bin is set to 210 GeV (950 GeV). These terms smooth the observed distributions using nearby bins, penalizing bins farther in the $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$ phase space. σ_{met} and σ_{meff} are width parameters. These parameters are both set to 20 GeV, which is a lower estimate of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} resolutions. As the amount of smoothing performed increases with the width parameters, a lower estimate corresponds to conservative smoothing. The same procedure is applied to down variations as well. The effect of the smearing on the first JER component for the 2016 200 GeV signal is shown in Figure 11.5.

The JMS is calculated by comparing the measurements of jet masses from the calorimeter and from the tracker [95]. 8 nuisance parameters are used to model these effects. These



(a) Without smoothing



(b) Smoothed

Figure 11.5: JER 1 systematics for the mc16a 200 GeV mass point beforeand after JER smoothing is applied. The width parameters σ_{met} and σ_{meff} are set to 20 GeV for the smoothing. The shaded blue region shows the statistical uncertainties on the nominal estimate.

include components to account for the differences between data and MC simulations, the differences between various MC generators, the statistical uncertainty, and the uncertainties arising from the tracks used in the calculation. In this analysis, the MC jet masses are only used in calculating the Higgs candidate masses. As these are dominated by the energy terms, the impact of JMS uncertainties is negligible.

The nominal value of $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated using a track-based soft term (TST) method that combines calorimeter-based hard objects with soft track terms. Soft terms are defined as deposited $E_{\mathrm{T}}^{\mathrm{miss}}$ not associated with any reconstructed object. The track-based method provides greater resiliency against pile-up than calorimeter-based methods. The differences between different MC generators, as well as those arising from changes to the detector and running settings, are set as systematics [89].

The effect of select signal systematics for 2016 130 GeV signal MC are shown in Figures 11.6 and 11.7. A total of 86 systematics were tested for each mass point for each year. These plots show the variations in the signal regions before the fit is performed. Figure 11.6 shows the first 4 components of the b-tagging efficiency systematic, the first components of the c-quark and light mis-tag systematics, the b-tagging extrapolation systematics, the JVT systematic, and the pile-up systematic. All flavor tagging systematics include the offline and online components. Figure 11.7 shows the first three components of the JER systematic, the JER data vs. MC systematic, 6 JES or other jet systematics, and 2 of the $E_{\mathrm{T}}^{\mathrm{miss}}$ systematics.

11.3 Theoretical Uncertainties

Monte Carlo samples are affected by uncertainties on our theory models. In this analysis, these are only applicable to our signal samples. Three types of uncertainties are considered.

First, there is the uncertainty on the cross section of higgsino production [75–77]. This is accounted for by scaling the limits up and down to the $\pm 1\sigma$ values of the cross section.

Second, there are uncertainties due to scale variations. These include the QCD factor-

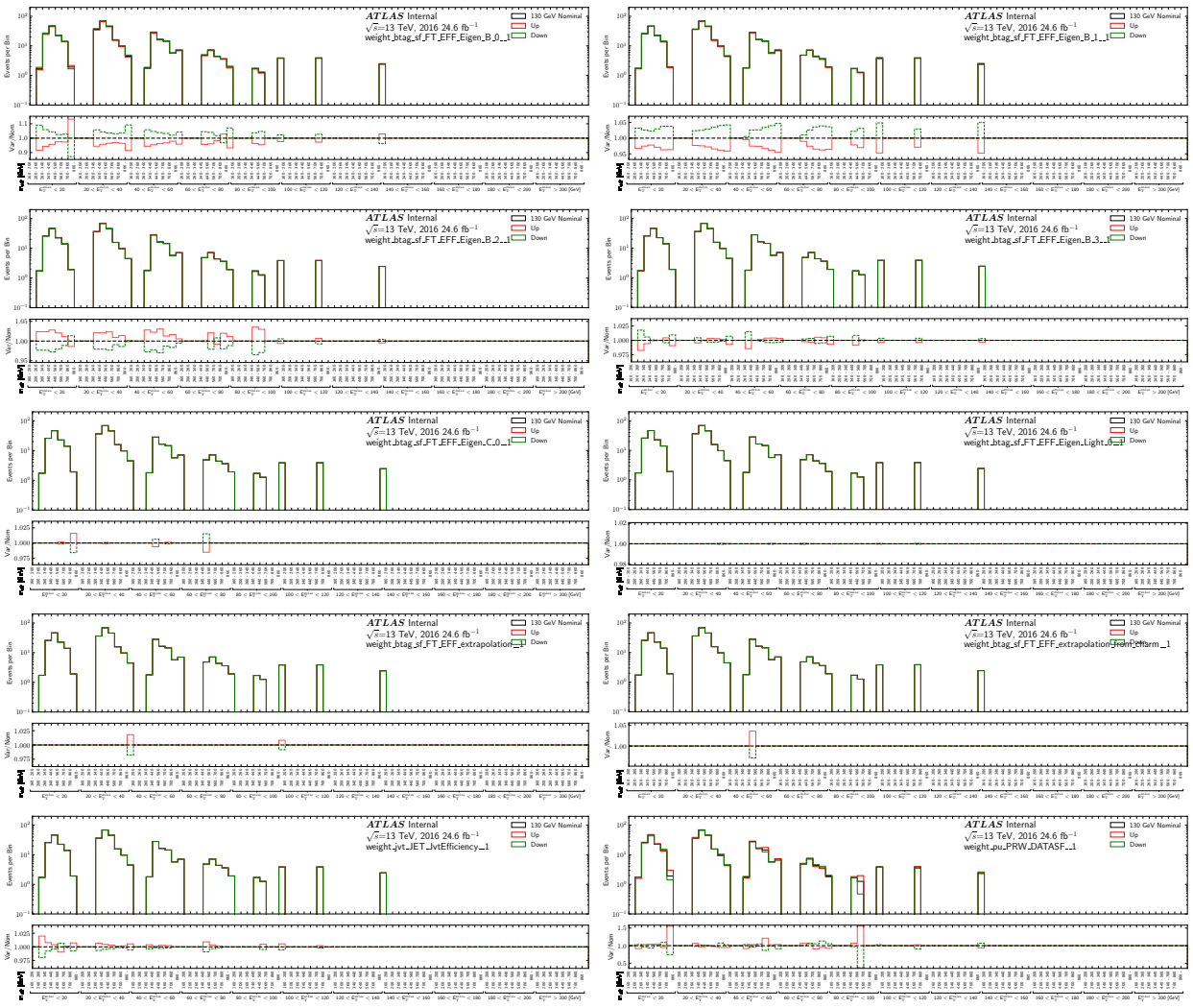


Figure 11.6: Pre-fit effect of selected signal systematics on the 2016 130 GeV signal MC samples. These systematics are calculated by applying weights to each event.

Figure 11.7: Pre-fit effect of selected signal systematics on the 2016 130 GeV signal MC samples. The JER systematics are shown before smoothing is applied. These systematics are calculated by reprocessing with different variations.

ization and renormalization scales as well as the CKKW merging scale. Finally, we have uncertainties on the tuning of the parton showering. To estimate the effect of these variations, alternative signal MC samples are generated using the scale variations, the merging variations, and the parton shower variations. Along with the nominal samples, these are processed through a SimpleAnalysis [96] framework to derive the expected number of truth events in each kinematic region and bin for the 130, 150, 200, 500, and 1100 GeV mass points. More truth points were generated at low masses as the $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$ distribution changes more rapidly in this regime. 1 million events were generated for each of the points except 1100 GeV, which used 600,000 events. The large number of bins in this analysis can cause large artificial fluctuations due to the statistical uncertainties. Each variation is therefore run through the same smoothing algorithm used for the JER uncertainties. The ratio of the varied to nominal events in each bin of the truth sample is used to derive a percent uncertainty. If any bins have zero predicted nominal truth events but non-zero events in a varied sample, the uncertainty for that bin is set to 100%. The uncertainties are then applied to the reconstructed signal samples as correlated shape systematics using the closest available mass point.

Chapter 12

SIGNAL CONTAMINATION

In order for the background estimation strategy to be reliable, the 2b CR, 4b CR, and 2b SR must have little or no signal compared to the 4b SR. To see why this is so, consider the unphysical but illustrative example where there is no background. In this scenario, each region is populated purely by signal events. The ABCD method's prediction from Chapter 9 would then give an estimate of the number of signal events in the 4b SR. Thus, even if a signal were present, no excess would be seen. More generally, if the signal contamination were high, the ABCD method would yield an estimate of $N^{4b}_{\mathrm{SR,S+B}}$ instead of the desired $N^{4b}_{\mathrm{SR,B}}$. A similar argument holds for the validation region: if the signal contamination in the VR was high, the systematics assigned in Chapter 11 would cover the difference between $N^{4b}_{\mathrm{SR,S+B}}$ and $N^{4b}_{\mathrm{SR,B}}$, eliminating our sensitivity to new signals. We thus require the ratio $N_{\mathrm{sig}}/N_{\mathrm{bkg}}$ to be small outside of the 4b SR.

The signal contamination was measured for each signal model in each background region. The results for 130 GeV signals are shown in Table 12.1. Higher mass signals have lower contamination due to their lower theoretical cross sections. The worst contamination falls in the 2017 4b validation region, with an $N_{\mathrm{sig}}/N_{\mathrm{bkg}}$ of 2.6%. The contamination in the 4b control regions are roughly a factor of four lower, while the 2b regions have negligible signal. These assume the theoretical prediction for the cross sections and a $BR(\tilde{H} \to h\tilde{G})$ of 100%. The estimates are therefore conservative, as some of these cross sections have been excluded by prior results. In addition, only $hh \to b\bar{b}b\bar{b}$ events are considered. Other hh decays are found to have negligible contributions due to their low branching ratios and the b-jet selections.

While the overall signal contamination is low for the CRs, VRs, and 2b SRs, the bin-wise signal contamination could still play an important role in the tails of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} distributions. This contamination makes the 4b sample look more signal-like, which could cause the reweighting procedure to learn the signal. This would increase the background

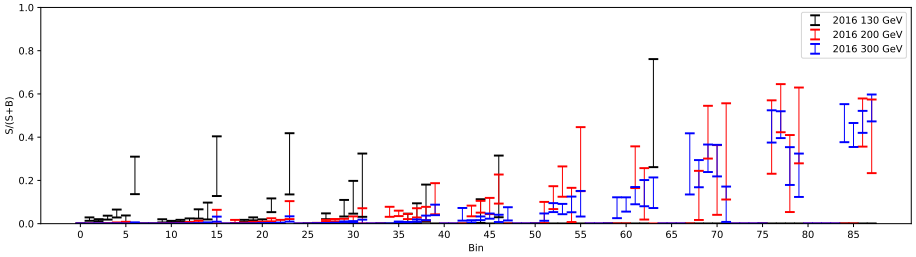
	CR	VR	SR
2016 2b	$4.1*10^{-5}$	$9.7*10^{-5}$	$8.0*10^{-4}$
2016 4b	$4.9*10^{-3}$	$2.1*10^{-2}$	$5.5*10^{-2}$
2017 2b	$6.6*10^{-5}$	$9.3*10^{-5}$	$1.1*10^{-3}$
2017 4b	$5.4*10^{-3}$	$2.6*10^{-2}$	$6.2*10^{-2}$
2018 2b	$5.6*10^{-5}$	$8.8*10^{-5}$	$8.8*10^{-4}$
2018 4b	$4.8*10^{-3}$	$2.0*10^{-2}$	$6.0*10^{-2}$

Table 12.1: S/B signal contamination in each region for the 130 GeV mass point. The background in the 4b SR is estimated as $N^{2b}_{\text{SR}} * N^{4b}_{\text{CR}}/N^{2b}_{\text{CR}}$ to preserve blinding. Signals are assumed to have the theory cross section with a $BR(\tilde{H} \to h\tilde{G})$ of 100%. $N_{\text{sig}}/N_{\text{bkg}}$ decreases as the signal mass increases.

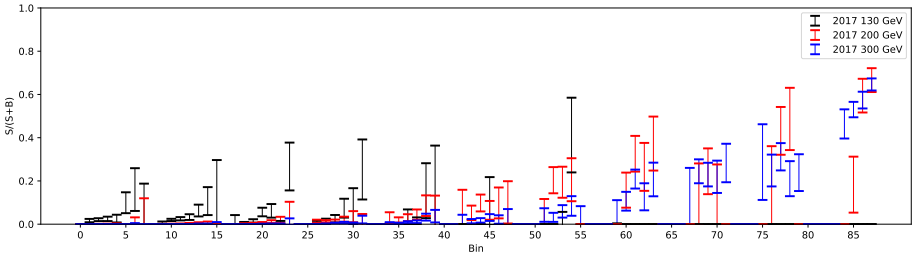
estimate, potentially hiding real signal. However, this effect is limited for two reasons. First, the BDT is not trained using any particular binning structure. Second, the BDT requires nodes to have at least 250 events, which should help wash out any effects from signal events in the tails.

The signal contamination for each bin in the 4b VR is shown in Figure 12.1. In addition to the signals (130, 200, and 300 GeV) shown in the plot, 150 and 250 GeV were also tested. The background estimate using reweighted 2b VR events is used in place of the 4b VR data events. This is done to preserve blinding, as any bin with a large expected signal contamination would otherwise give hints as to whether a signal is present. The maximum signal contamination occurs for the 2016 bin with $160 < E_{\text{T}}^{\text{miss}} < 180$ GeV and $m_{\text{eff}} > 860$ GeV. Here, $N_{\text{sig}}/(N_{\text{sig}} + N_{\text{bkg}})$ is 78%.

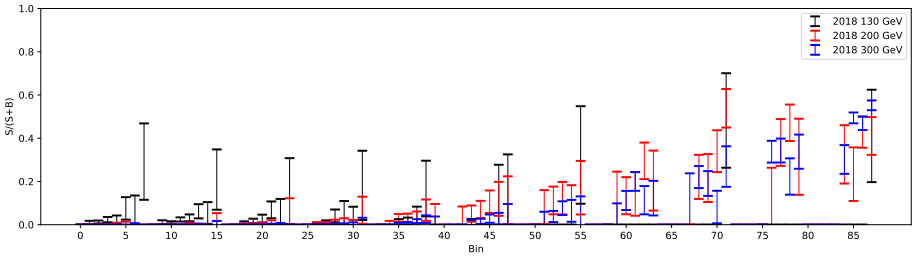
Given the large signal contamination in high $E_{\text{T}}^{\text{miss}}$ bins, the background estimation procedure was tested for robustness to signal contamination. The BDT reweighting was repeated with signals of masses $m_{\tilde{H}}$=130 GeV, 150 GeV, 200 GeV, 250 GeV, and 300 GeV injected into the 2b CR, VRs, and SRs as well as the 4b CRs and VRs. The results were first checked in the discovery regions, as shown in Table 12.2. The signals were scaled down to the highest cross sections not excluded by previous ATLAS [26] and CMS [97] results. The signals were



(a) 2016



(b) 2017



(c) 2018

Figure 12.1: Bin-wise $N_{\text{sig}}/(N_{\text{sig}} + N_{\text{bkg}})$ signal contamination for the 4b VR. The signals assume the theory cross section with a $BR(\tilde{H} \to \tilde{h}G)$ of 100%. The uncertainties only include the component from signal MC statistics.

	150 GeV	300 GeV
Nominal	8.62σ	8.968σ
130 injection	8.60σ	8.61σ
150 injection	8.59σ	8.969σ
200 injection	8.60σ	8.82σ
250 injection	8.60σ	8.67σ
300 injection	8.60σ	8.71σ

Table 12.2: Expected significance of discoveries with and without signal injections. Only statistical uncertainties are included.

further scaled down to the lowest cross section that would still yield a 5σ discovery, using an estimate of the discovery region systematics. This only affected the 200 and 250 GeV mass points. The final scale values were 0.18, 0.4, 0.75, 0.6, and 0.45 for the 130, 150, 200, 250, and 300 GeV signals respectively. The effect of the injected signals was small, slightly reducing the expected significance in the discovery regions. This is expected as the discovery regions use large bins, making them less sensitive to the tails of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} distributions. These studies were done using preliminary definitions of the discovery regions and their systematics and are thus not directly comparable to Chapter 10.

Initial studies without bootstraps showed the 130 GeV and 200 GeV signal injections had the greatest effects on the limits between 130 and 300 GeV. These points were investigated further using 10 bootstraps. The expected limits set the observations equal to the background prediction while the observed limits were set using observations equal to the background prediction plus the scaled signals. Results are shown in Figure 12.2. The expected limits were up to 3.4% better (using the 130 GeV injection and evaluating at the 130 GeV mass point) and up to 10.8% worse (using the 200 GeV injection and evaluating at the 300 GeV mass point) than the nominal limits. The observed limits had much smaller variations, improving by up to 1.3% and worsening by up to 1.7%.

From this, it was concluded that no special care is needed to mitigate signal contam-

ination. The 130 GeV injection has minimal effects on the limits. Though the expected limits are worse in some cases for the 200 GeV injection, the effects on the observed limits (and therefore the final results) are minimal. In addition, the size of the effect should scale with signal strength. If we set limits close to our sensitivity (≈ 0.2 times signal strength at 200 GeV), the signal will have at most 1/3 the strength used in the injection tests and the potential impact from signal contamination will be reduced accordingly.

It is important to note that any effects from signal contamination are already present in our data and reflected in our results shown in Chapter 13. The injection tests discussed here show the effect of adding additional signal, under the assumption that the effects scale with signal strength. As injecting signals makes our limits worse under most scenarios, this means the results shown are slightly conservative with respect to the results that could be hypothetically obtained without the presence of signal contamination.

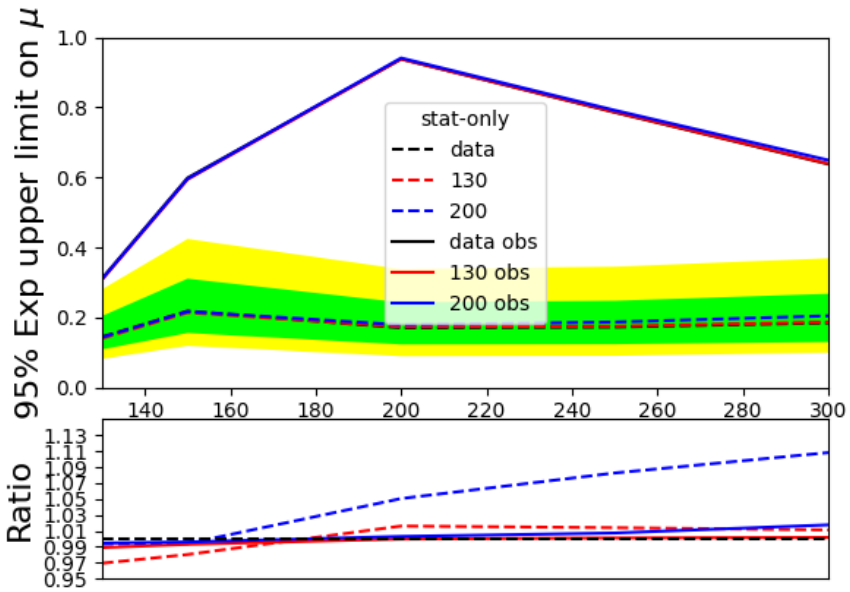


Figure 12.2: Limits with injected 130 GeV and 200 GeV signals. Signals are scaled to the smaller of the highest non-excluded cross section and the lowest signal strength that would still result in a 5σ discovery, using an estimate of the systematics. To simulate the observed limits without unblinding, the observations are set to the background prediction plus the scaled signals.

Chapter 13

RESULTS

13.1 Signal Region Data

The observations and their corresponding background estimates for each bin are shown in Figure 13.1. The distributions for 250 GeV higgsinos are shown for comparison. These plots use pre-fit values of the fit parameters. The uncertainties and significances shown are statistical only.

Given the large number of bins in this analysis, some deviations from the background predictions are expected due to statistical variations. In particular, 32% of measurements are expected to fall outside of the 1σ band, 4.55% are expected to fall outside of the 2σ band, and 0.27% are expected to fall outside of the 3σ band. If the discrepancies between background and observations are purely due to statistical effects, we would expect these deviations to form a Gaussian distribution of mean 0 and standard deviation 1. The actual distributions and their fits to Gaussian are shown in Figure 13.2.

The 2016 channel is nearly consistent with statistical expectations. The distribution of the deviations has a high probability of being described by a Gaussian (78%). The mean of this Gaussian is nearly consistent with 0, but slightly biased towards negative values (meaning the predicted background is larger than the observations). We would expect roughly 22 bins to fall outside of the 1σ band and 3 bins to fall outside of the 2σ band. These are quite close to the actual numbers of 18 and 1 respectively.

The 2017 channel is less consistent with statistical expectations. The distribution of the deviations is less compatible with a Gaussian and is significantly biased towards negative pulls (higher background than observation). For this channel, we would expect roughly 22 bins to fall outside of the 1σ band and 3 bins to fall outside of the 2σ band due to statistical variations. We observe notably higher deviations, with 24 and 8 bins falling outside of the 1 and 2σ ranges respectively. Some of this could be due to the normalization uncertainty,

which is not included in the significance calculation as it must be applied coherently across bins. In addition, bin migrations could be playing a role in the 160 GeV$< E_{\mathrm{T}}^{\mathrm{miss}} <$ 180 GeV bin, where adjacent m_{eff} bins have $> 2\sigma$ and $< -2\sigma$ deviations from the background model.

The 2018 channel is moderately consistent with statistically distributed deviations. The distribution is consistent with a Gaussian of mean 0 and standard deviation 1. Approximately 22 and 3 bins are expected to fall outside of the 1 and 2σ bands respectively, in good agreement with the observed 20 and 3 bins.

The largest excess occurs in the 2017 channel, this time in the $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV, $m_{\mathrm{eff}} >$ 860 GeV bin. This has a background of 1.51 ± 0.35 events, an observation of 6 events, and a discrepancy of 2.58σ. Notably, no bins differ from the background estimate by more than 3σ. Given the number of bins, we expect 0.6 events to fall outside of this range due to statistical effects.

The observations are also checked with respect to $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} individually. These results are shown in Figure 13.3. Good agreement is observed up to $\approx 150 - 200$ GeV in $E_{\mathrm{T}}^{\mathrm{miss}}$ and 1 TeV in m_{eff}. The agreement is difficult to judge for higher $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} values due to limited statistics.

13.2 Statistical Analysis

Given a background model, signal model, and set of systematic uncertainties, we need to determine the probability of new physics being observed. In addition, if there is no statistically significant excess, we need to set upper limits on the possible cross sections of new physics processes.

First, a likelihood function is constructed in the discovery regions, giving the probability that the observations can be described using the model. This likelihood, represented as $L(\mu, \boldsymbol{\theta})$, is a function of the signal strength μ (normalized such that the signal theory predicts $\mu = 1$) and the nuisance parameters (NPs) $\boldsymbol{\theta}$. Normalization and luminosity systematics

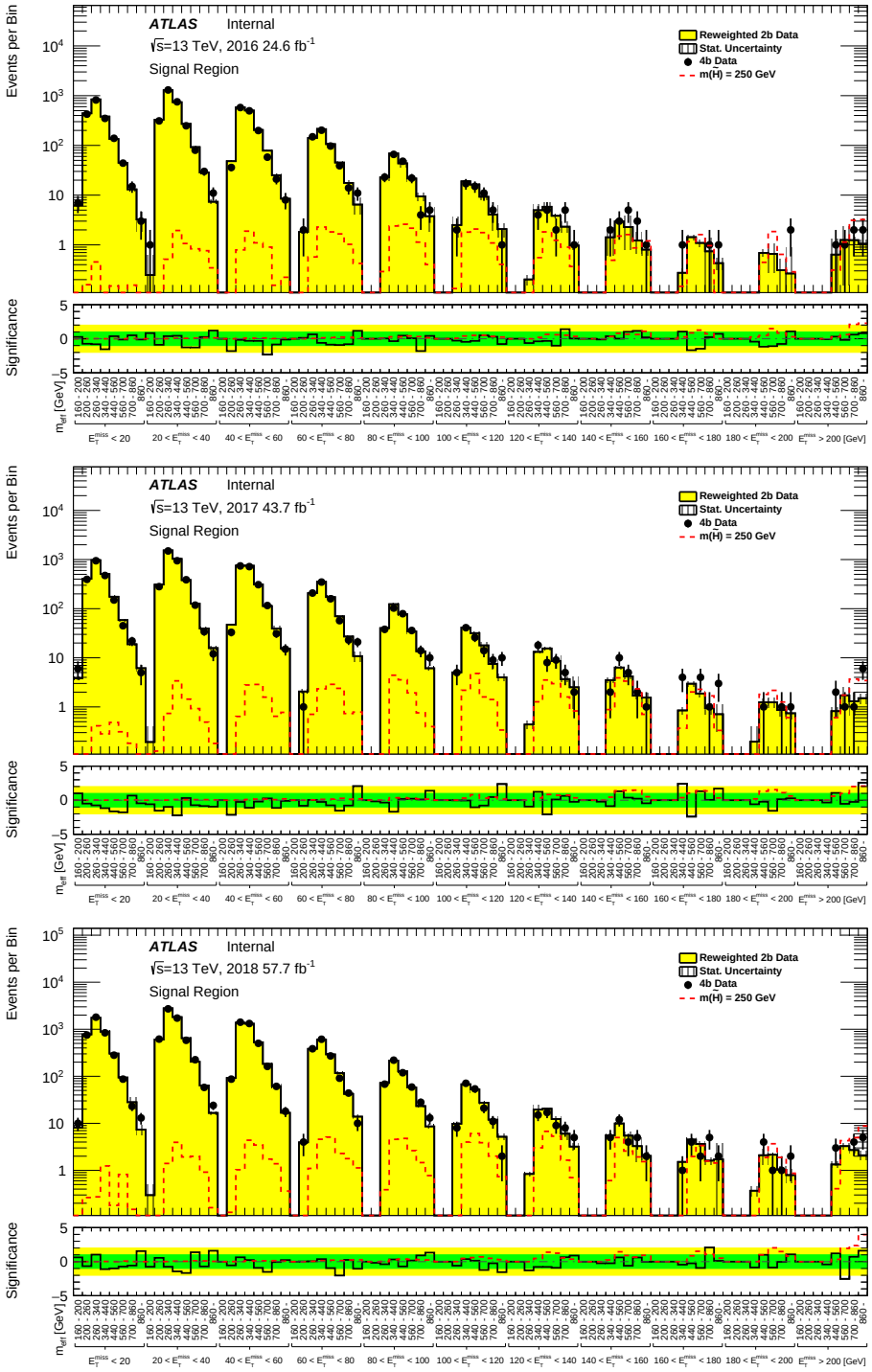
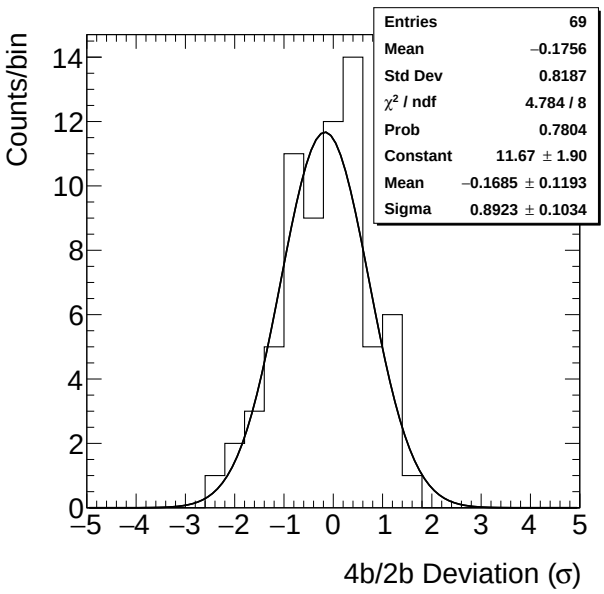
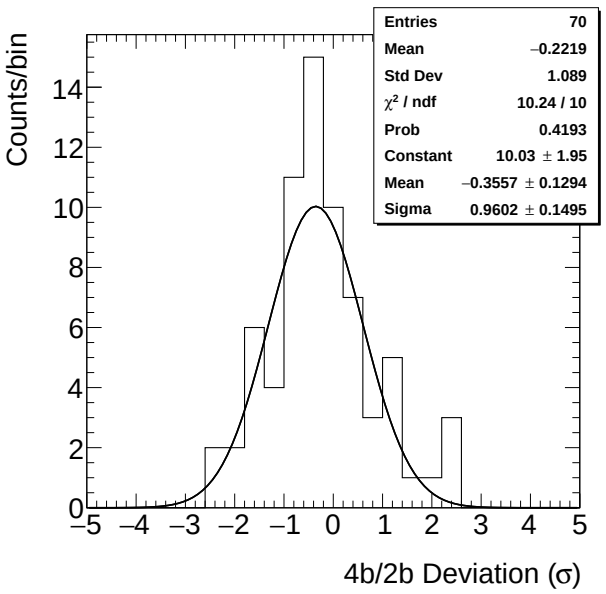


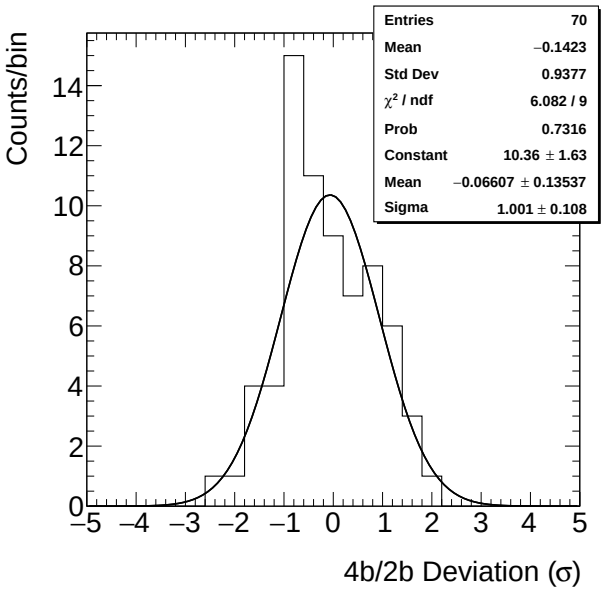
Figure 13.1: Differences between the 4b and reweighted 2b distributions in the 2016, 2017, and 2018 SRs. These use the pre-fit values of the background estimate. Signals for $m_{\tilde{\mathrm{H}}} = 250$ GeV are shown for comparison. The non-closure and transfer shape uncertainties are included. The significance of the signal is defined as the significance that would be observed if the number of observed 4b events were equal to the number of expected signal events plus the estimated background events.



(a) 2016

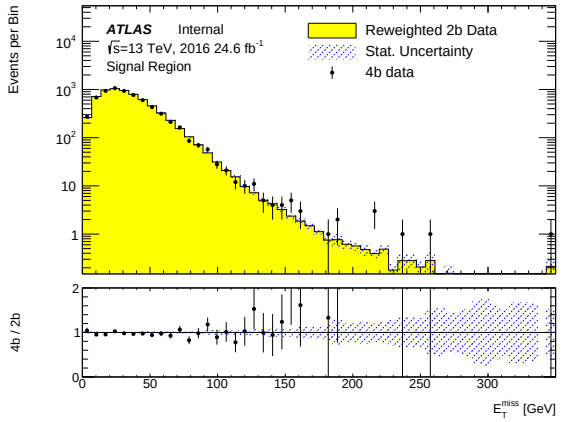


(b) 2017

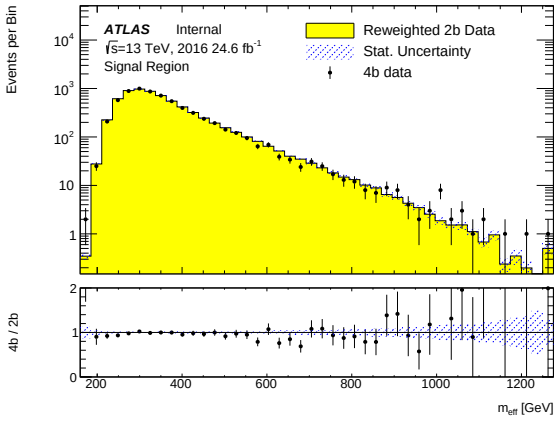


(c) 2018

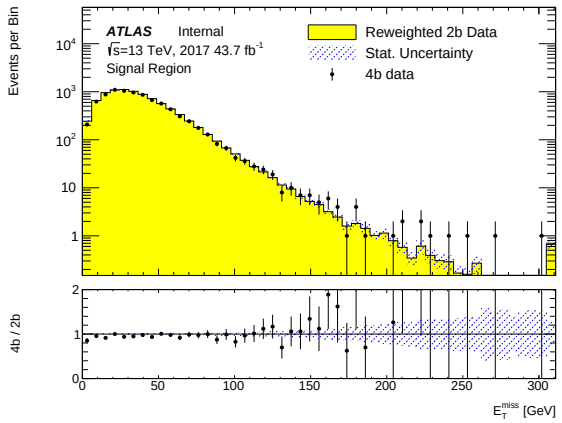
Figure 13.2: Distributions of the differences between the 4b and reweighted 2b distributions in the 2016, 2017, and 2018 SRs. The distributions are fit to Gaussians. Postive σ corresponds to the case where more 4b events are observed than reweighted 2b events. These use the pre-fit values of the background estimate. The non-closure and transfer shape uncertainties are included. Bins with no events in the 2b or 4b samples are excluded.



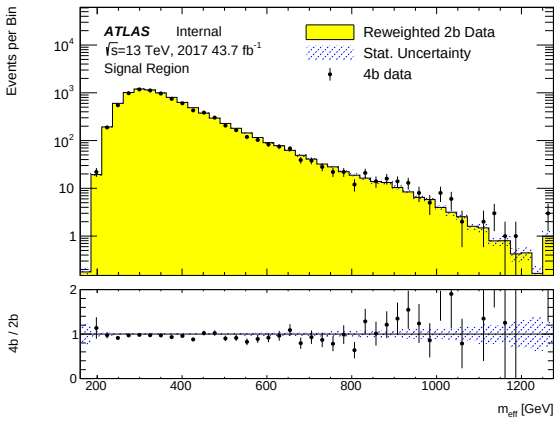
(a) 2016 $E_{\mathrm{T}}^{\mathrm{miss}}$



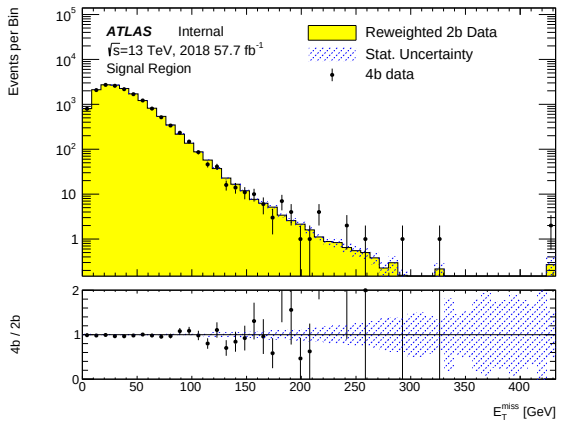
(b) 2016 m_{eff}



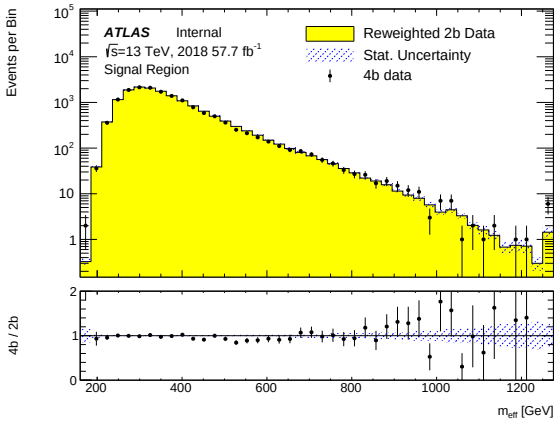
(c) 2017 $E_{\mathrm{T}}^{\mathrm{miss}}$



(d) 2017 m_{eff}



(e) 2018 $E_{\mathrm{T}}^{\mathrm{miss}}$



(f) 2018 m_{eff}

Figure 13.3: Comparisons of 4b and reweighted 2b binning variables in the signal regions. The last bin is an overflow bin. The blue shading shows the uncertainty due to the reweighting and 2b statistics.

make use of Gaussian constraints while shape systematic and statistical NPs make use of Poisson constraints. This is used to construct the discovery test statistic q_0, defined as

$$q_0 = \begin{cases} -2ln\frac{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{13.1}$$

where $\hat{\mu}$ is the best fit value of μ, $\hat{\boldsymbol{\theta}}$ is the best fit value of $\boldsymbol{\theta}$, and $\hat{\hat{\boldsymbol{\theta}}}$ is the best fit value of $\boldsymbol{\theta}$ with the fit constrained to a fixed value of μ.

Next, the CL_{s+b} is defined as the probability p_s for the signal plus background model to yield observations that have a less signal-like test statistic than the observations. For q_0, this is equivalent to the p-value of the background-only fit. A p-value is considered statistically significant if it corresponds to a 5σ deviation or higher [98–100].

In the absence of a statistically significant observation, upper limits are set at a 95% confidence level (CL). This is performed using the model-dependent exclusion regions and uses a more complex test statistic given by [100]

$$\tilde{q}_\mu = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \\ -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{13.2}$$

In addition to CL_{s+b}, CL_b is defined as the probability for the background-only model to yield predictions that have a less signal-like test statistic than the observations [98–100]. The ratio

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{p_s}{1-p_b} \tag{13.3}$$

is then calculated. Limits are set at the μ values corresponding to $CL_s(\mu) = 0.05$ [98, 99].

Signal channel	N_{obs}	N_{pred}	$\langle\epsilon\sigma\rangle^{95}_{\text{obs}}$ [fb]	S^{95}_{obs}	S^{95}_{exp}	p_0	Z
SR-150	1790	1859.6±49.6	0.73	93	127^{+48}_{-34}	0.5	0.00
SR-300	97	77.0±5.3	0.32	39	22^{+9}_{-6}	0.03	1.83

Table 13.1: Results for the discovery regions. The columns show the number of observed events (N_{obs}), the number of predicted events (N_{pred}), the 95% CL limits on the visible cross section, the 95% CL limits on the number of signal events, the expected limits on the number of signal events, and the discovery p-value and significance. p-values are capped at 0.5. Results are obtained using 20,000 pseudo-experiments. The limits on the cross section are obtained by dividing the limits on the number of signal events by the integrated luminosity of 126 fb^{-1}.

The statistical framework is implemented through pyhf [101, 102] using the MINUIT optimizer [103].

13.3 Model-Independent Results

In the discovery regions, limits are set using 20,000 pseudoexperiments. The results are shown in Table 13.1 and the post-fit yields are shown in Table 13.2. A slight deficit of $\approx 1\sigma$ was observed in the low-$E_{\text{T}}^{\text{miss}}$ discovery region (SR-150, optimized for 150 GeV higgsinos). The high-$E_{\text{T}}^{\text{miss}}$ discovery region (SR-300, optimized for 300 GeV higgsinos), on the other hand, yielded a slight excess of 1.83σ. For the $E_{\text{T}}^{\text{miss}} > 20$, $m_{\text{eff}} > 560$ GeV region of phase space, the limit on the cross section times efficiency of Beyond the Standard Model physics is set at 0.73 fb. Though the definition of the region is slightly different, this is roughly 3 times stronger than the limits set on the low-$E_{\text{T}}^{\text{miss}}$ region ($E_{\text{T}}^{\text{miss}} > 0$, $m_{\text{eff}} > 440$ GeV) in the previous iteration of this analysis [26]. The high-$E_{\text{T}}^{\text{miss}}$ region is slightly looser than in the previous version, only differing by having a m_{eff} threshold of 340 GeV instead of 440 GeV. Despite this and the slight excess, the limit on the efficiency times cross section of new physics improved by a factor of 2.8 to 0.32 fb.

	Background
SR-150	1819.7±32.2
SR-300	82.3±4.7

Table 13.2: Post-fit background yields for the 150 GeV and 300 GeV discovery region background-only fits. Each region consists of a single bin.

	Background	Total	Obs
2016	6096.7±69.5	6096.7±69.5	6066
2017	8179.3±88.0	8179.3±88.0	8154
2018	14366.4±113.4	14366.4±113.4	14334

Table 13.3: Post-fit yields for the 130 GeV signal. Results are obtained by summing over all bins. The total yields are signal plus background.

13.4 Model-Dependent Results

As no significant excess was seen in the discovery regions, limits are drawn for the higgsino model under consideration. The full $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} distributions from Figure 13.1 are used, except that for each signal, bins with no signal MC events are excluded from the fit. Each year of data is treated as a separate analysis category and all background, signal, signal theory, and luminosity systematics are included. While the analysis is optimized to target the $hh \to b\bar{b}b\bar{b}$ final state, all hh decays are included in the signal model. The fits are performed using the asymptotic approximation [100]. The post-fit yields are shown in Tables 13.3-13.7.

Limits are set on the cross section of higgsino pair production assuming a $\tilde{H} \to h\tilde{G}$ branching ratio of 100%. These are shown in Figure 13.4, with the limits for the previous iteration of the analysis overlaid in Figure 13.5 for comparison. Higgsinos with masses of up

	Background	Total	Obs
2016	5933.2±69.1	5933.2±69.1	5903
2017	8440.1±90.5	8440.1±90.5	8414
2018	15302.5±117.8	15302.5±117.8	15273

Table 13.4: Post-fit yields for the 150 GeV signal. Results are obtained by summing over all bins. The total yields are signal plus background.

	Background	Total	Obs
2016	6322.9±70.0	6341.7±69.9	6304
2017	8067.8±86.0	8099.5±85.9	8074
2018	15400.0±113.6	15445.5±113.5	15408

Table 13.5: Post-fit yields for the 200 GeV signal. Results are obtained by summing over all bins. The total yields are signal plus background.

	Background	Total	Obs
2016	5967.2±67.3	5976.8±67.3	5951
2017	8098.8±85.9	8114.0±85.9	8088
2018	14897.9±111.3	14920.2±111.3	14882

Table 13.6: Post-fit yields for the 250 GeV signal. Results are obtained by summing over all bins. The total yields are signal plus background.

to 530 GeV (740 GeV) expected are excluded at the 95% confidence level. This is a significant improvement over the previous iteration of the analysis. The excluded cross sections improve by a factor of 1.2-3.5, depending on the mass point [26]. If we simply scaled the luminosity of the previous analysis, we would expect to improve by roughly a factor of 2.3. As the 130 and 150 GeV mass points were already strongly excluded, the optimization for this analysis focused more on the 200-300 GeV regime. In this phase space, the limits improved by a factor of 2.3-3.5. In addition, this analysis excluded the 250 GeV mass point, which the previous analysis was unable to do.

Results are also interpreted as limits on $BR(\tilde{H} \to h\tilde{G})$ by instead assuming the theoretical cross section for higgsino production. All higgsinos are assumed to decay as either $\tilde{H} \to h\tilde{G}$ or $\tilde{H} \to Z\tilde{G}$. The limits for this interpretation are shown in Figure 13.6. The hZ and ZZ decays peak outside of our signal regions and would thus be seen as background rather than signal. Therefore, the limit on the branching ratio is equivalent to the square root of the

	Background	Total	Obs
2016	5944.8±67.2	5951.9±67.2	5923
2017	7856.2±84.5	7866.7±84.5	7841
2018	12271.5±100.2	12287.6±100.2	12236

Table 13.7: Post-fit yields for the 300 GeV signal. Results are obtained by summing over all bins. The total yields are signal plus background.

limit on $\sigma/\sigma_{\text{theory}}$ shown in the underlay of Figure 13.4.

We observe a slight excess for higgsino masses of 200 GeV and higher. This excess grows to $> 2\sigma$ for higgsinos with $m_{\tilde{H}} > 400$ GeV. For signals with such high masses, nearly all events fall in the $E_{\text{T}}^{\text{miss}} > 200$ GeV, $m_{\text{eff}} > 860$ GeV bin. In particular, this means the signals look nearly identical (up to cross section scaling) for large higgsino masses. Therefore, any excess seen by one such higgsino mass point is expected to occur for the others as well.

The kinematics of high-mass higgsinos also help explain the excess in the limits. From Figure 13.1, we see that there is an excess in the $E_{\text{T}}^{\text{miss}} > 200$ GeV, $m_{\text{eff}} > 860$ GeV bin for each channel. This excess is 0.8σ for 2016, 2.6σ for 2017, and 1.6σ for 2018. The excess in the limits is therefore mostly driven by the 2017 excess.

There are three potential sources of excesses. These are mismodeling of the background estimate, statistical fluctuations, and new physics. The excess is not believed to be due to mismodeling, given no such excess is seen in the validation region (Figure 11.2). In addition, mismodeling would be expected to affect adjacent bins, which do not see an excess.

Neither is the excess likely explained by new physics. If this were the case, the significance of the excess should grow with luminosity. In particular, this means the 2018 excess should be larger than the 2017 excess. However, the observed excess is much larger in 2017 than 2018. While it possible that statistical fluctuations could combine with a real signal to give these observations, this still supplies evidence against the new physics interpretation.

Finally, it is possible the excess is due entirely to statistical variations. As discussed above, this analysis has enough bins that we expect several of them to have excesses greater than 2σ. In addition, the number of such bins we observe (13) is not too much higher than the number expected (9), even before considering that some of them could be explained by our background systematics. If the 2017 excess is indeed just a statistical fluctuation, it simply fell in an unlucky spot for our sensitivity to high-mass higgsinos. As most events for these signals would fall in the $E_{\text{T}}^{\text{miss}} > 200$ GeV, $m_{\text{eff}} > 860$ GeV bin, any statistical fluctuation there would have a large impact on their limits.

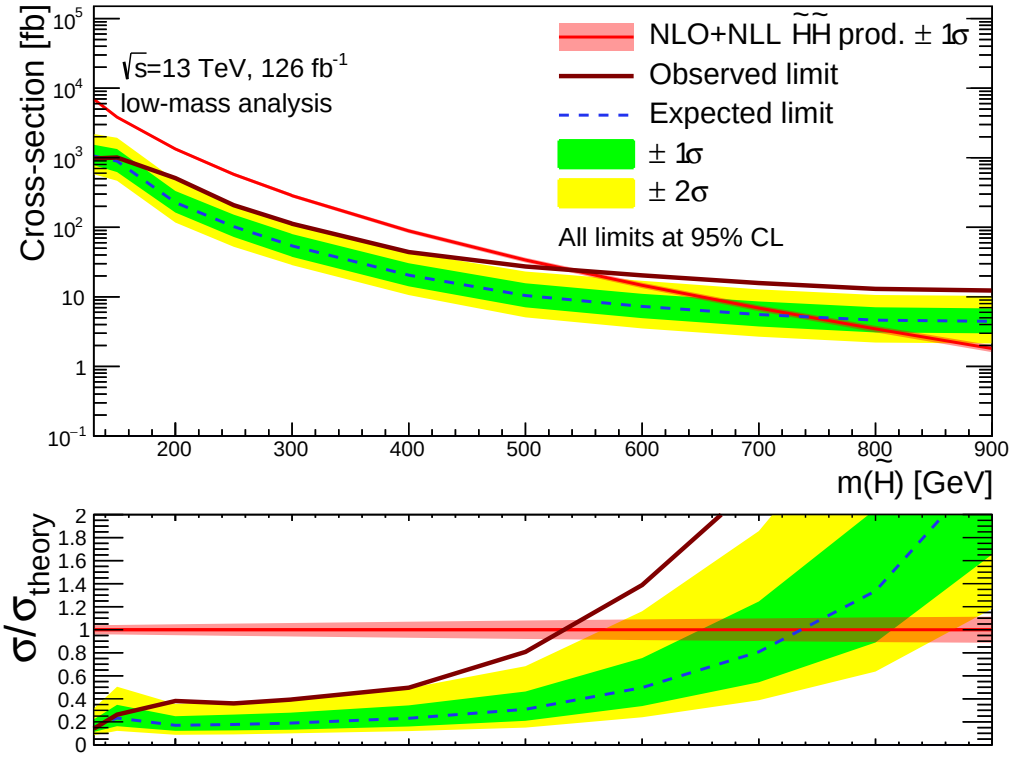


Figure 13.4: 95% CL limits on higgsino production. This assumes a $\tilde{H} \to h\tilde{G}$ branching ratio of 100%.

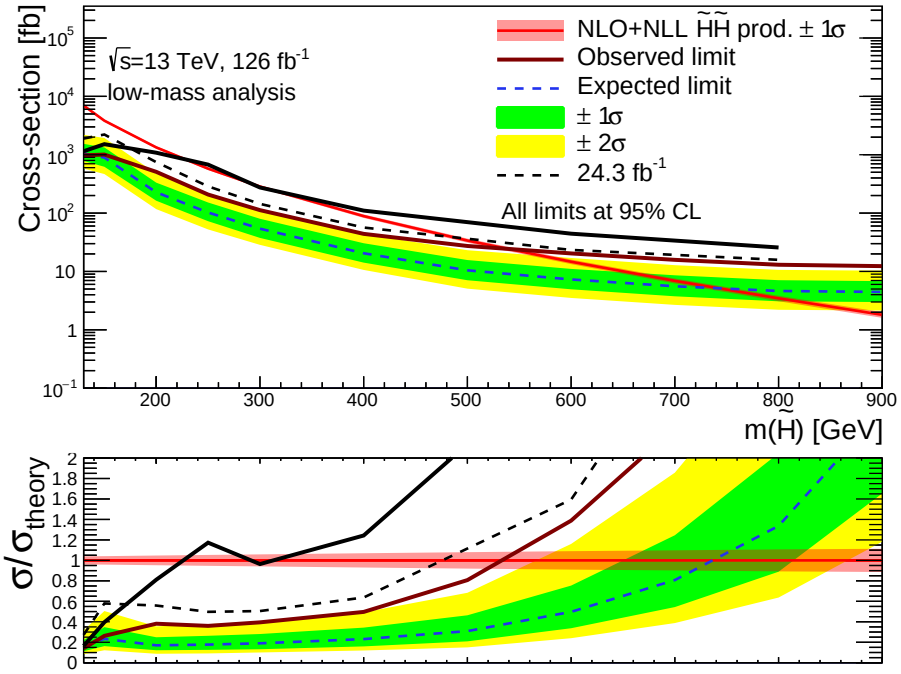


Figure 13.5: 95% CL limits on higgsino production. This assumes a $\tilde{H} \to h\tilde{G}$ branching ratio of 100%. The black dashed (solid) line shows the expected (observed) limits from the previous analysis with 24.3 fb^{-1}.

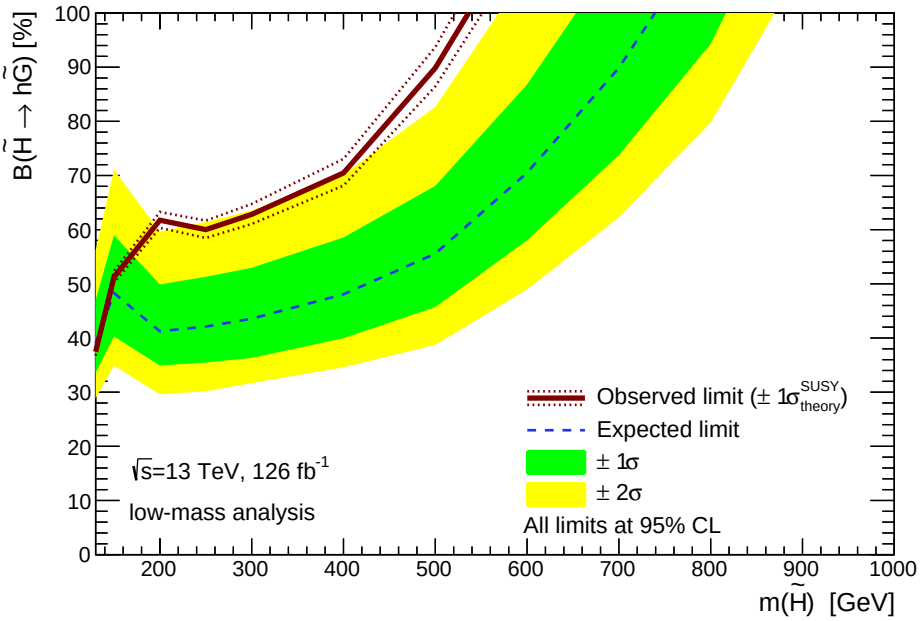


Figure 13.6: Limits on the branching ratio for low-mass analysis. This assumes the theory cross section, and that all higgsinos decay to either a Higgs boson or a Z boson.

Chapter 14

CONCLUSION AND FUTURE IDEAS

Supsersymmetry is an elegant, well-motivated theory that would neatly solve several outstanding problems in phyiscs. Unfortunately, the Large Hadron Collider has, as of yet, seen no definitive sign of it. This makes higgsino searches a tantalizing prospect. Searches for high-mass superparticles have all turned up empty, and the simplest SUSY models require higgsinos to be relatively light. This motivates more thorough searches in the low-mass regime.

This thesis presents such a search. Starting with the prior analysis as a baseline, several key improvements were made. The triggers were re-optimized to provide maximal sensitivity and accurate calibrations. The binning was made significantly finer to better distinguish signals from the large QCD multijet background. These optimizations paid special focus to the key $m_{\tilde{H}} = 200 - 300$ GeV phase space where an excess was previously observed. The background estimation strategy, methods of constructing Higgs boson candidates, and estimates of the systematic uncertainties were all made more robust to help avoid the possibility of a spurious excess. The integrated luminosity was increased from 24.3 fb^{-1} to 126 fb^{-1}.

Two discovery regions were defined, targeting signals with $m_{\tilde{H}} = 150$ GeV and $m_{\tilde{H}} = 300$ GeV respectively. No significant excess was observed. For the 150 GeV region, the predicted background was 1859.6 ± 49.6 events, while 1790 events were observed. For the 300 GeV region, the expected background and actual observed number of events were 77.0 ± 5.3 and 97 respectively. Given the lack of a discovery, limits were drawn using the GGM/GMSB model. Assuming a $\tilde{H} \to h\tilde{G}$ branching ratio of 100%, higgsinos with masses up to 530 GeV (740 GeV expected) were excluded, with limits on the cross section as low as 14% (15% expected) of the theory value. Alternatively, assuming the theory cross section, the $\tilde{H} \to h\tilde{G}$ branching ratio for higgsino masses less than 300 GeV was constrained to $< 63\%$ ($< 48\%$ expected).

Despite all of the improvements over the previous analysis, no statistically significant

signs of supersymmetry were observed. Given these results, the remaining phase space for light higgsinos is narrowing. When combined with other analyses targeting leptonic or other final states, the GGM/GMSB model with gravitino LSP is ruled out for low-mass higgsinos with the predicted theoretical cross section [85–87].

This leaves limited options for continuing to pursue light higgsinos. The first option is the simplest: continue to create new techniques using the same general strategy, ruling out ever lower cross sections. Unless a strong motivation is found for this process to occur at these cross sections, this option is of limited utility. The second option is to critically examine the assumptions in our model. One such assumption stands out: we assume that the higgsinos decay promptly in our detector. Targeting this assumption would require a dedicated search and would provide useful constraints on the remaining possibilities. Finally, we could critically examine our model itself. Other models, including higgsino LSPs, are not as constrained, allowing further searches to provide useful results [86, 87].

Within this framework, there are many improvements that could be made for future searches. One idea that will be implemented for Run 3 is asymmetric b-jet trigger thresholds. The bulk of our sensitivity is driven by our 2b2j trigger, where each b-jet is required to pass the same p_{T} threshold. By using different thresholds for each b-jet, we could potentially capture more events while maintaining a low trigger rate. Looking further into the future, hardware tracking in the trigger could yield higher efficiencies at comparable rates.

Machine-learning techniques could also be utilized in more depth and breadth. The background reweighting could use stronger techniques, such as a neural network. The pairing of Higgs candidate jets could also be done with machine learning, as the current min-ΔR pairing has limited signal efficiency. Finally, machine learning could be employed to separate signal and background events. Some of these techniques have already been implemented or tested for the $hh \to b\bar{b}b\bar{b}$ analysis discussed in Chapter 16.

Finally, new regions and background reduction strategies could be employed. As the dominant background is QCD with mismeasured $E_{\mathrm{T}}^{\mathrm{miss}}$, a veto could be constructed using

the angle between the jets and missing energy for higher energy events. Potential new regions include 3b regions where one b-jet fails the tagging and boosted topologies with two large-radius jets. As we gather more data, other Higgs boson decay channels with less statistics but lower background, such as $hh \to bb\tau^+\tau^-$ or $hh \to bb\gamma\gamma$, could provide complimentary constraints.

Part IV

Di-Higgs Analyses

Chapter 15

κ_λ REWEIGHTING

In order to design and run analyses to constrain κ_λ, it is necessary to have simulations of ATLAS data for various κ_λ values. Given the range of values we would like to test, this would be computationally expensive. In addition, precise results would require fine spacing between the simulated κ_λ, raising the costs even further.

To solve this, a strategy was implemented to reweight SM samples ($\kappa_\lambda = 1$) to any value of κ_λ. This strategy follows the prescription in [27]. As described in Section 2.3, two diagrams contribute to ggF di-Higgs production at leading order, the box diagram and the triangle diagram. The box diagram is independent of κ_λ, while the triangle diagram's matrix element picks up a factor of κ_λ. This allows the differential cross section to be written as

$$\frac{d\sigma}{d\Phi} = A(\Phi) + B(\Phi)\kappa_\lambda + C(\Phi)\kappa_\lambda^2 \tag{15.1}$$

where Φ represents a region of phase space and A, B, and C are constants that only depend on Φ [27]. As we expect the acceptance times efficiency for di-Higgs analyses to only depend on the invariant mass m_{HH}, we select this to use for our variable Φ.

Equation 15.1 has three unknowns (A, B, and C). For each value of κ_λ we simulate, we obtain the cross section for one equation. Therefore, by simulating three values of κ_λ, we obtain a system of linear equations that can be solved exactly. In particular, choosing the κ_λ values of 0, 1, and 20, we obtain

$$A = (\frac{d\sigma}{dm_{HH}})_{(\kappa_\lambda=0)} \tag{15.2}$$

$$B = \frac{1}{380} * (-399(\frac{d\sigma}{dm_{HH}})_{(\kappa_\lambda=0)} + 400(\frac{d\sigma}{dm_{HH}})_{(\kappa_\lambda=1)} - (\frac{d\sigma}{dm_{HH}})_{(\kappa_\lambda=20)}) \tag{15.3}$$

$$C = \frac{1}{380} * (19(\frac{d\sigma}{dm_{HH}})_{(\kappa_\lambda=0)} - 20(\frac{d\sigma}{dm_{HH}})_{(\kappa_\lambda=1)} + (\frac{d\sigma}{dm_{HH}})_{(\kappa_\lambda=20)}) \tag{15.4}$$

To solve this system of equations, we begin by generating 10 million truth-level events for each of $\kappa_\lambda = 0$, 1, 10, and 20 using POWHEG BOX v2 [79–81]. The extra point at $\kappa_\lambda = 10$ is used for validation. The decays of the Higgs bosons are not taken into account, allowing the results to be applied to all decay channels. The normalized results of these histograms are shown in Figure 15.1.

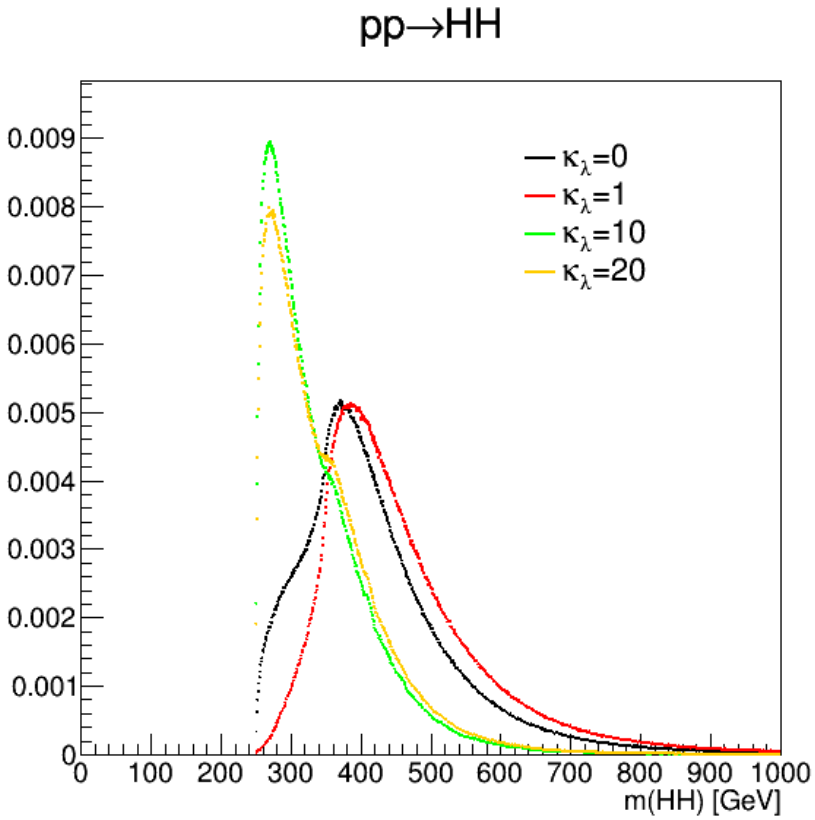


Figure 15.1: Normalized truth-level m_{HH} distributions for various κ_λ.

It can be seen that larger values of $|\kappa_\lambda|$ tend to yield lower values of m_{HH}. This is because the triangle diagram, which has an increasing contribution as $|\kappa_\lambda|$ increases, is peaked at lower values of m_{HH} as shown in Figure 15.2.

Using Equations 15.2 to 15.4, we can then generate the m_{HH} spectrum for any value of κ_λ. Figure 15.3 validates this procedure by showing the comparison between the $\kappa_\lambda = 10$ distribution from simulation vs. from the linear combination.

Using the m_{HH} distributions calculated above, weights are derived as the ratios between

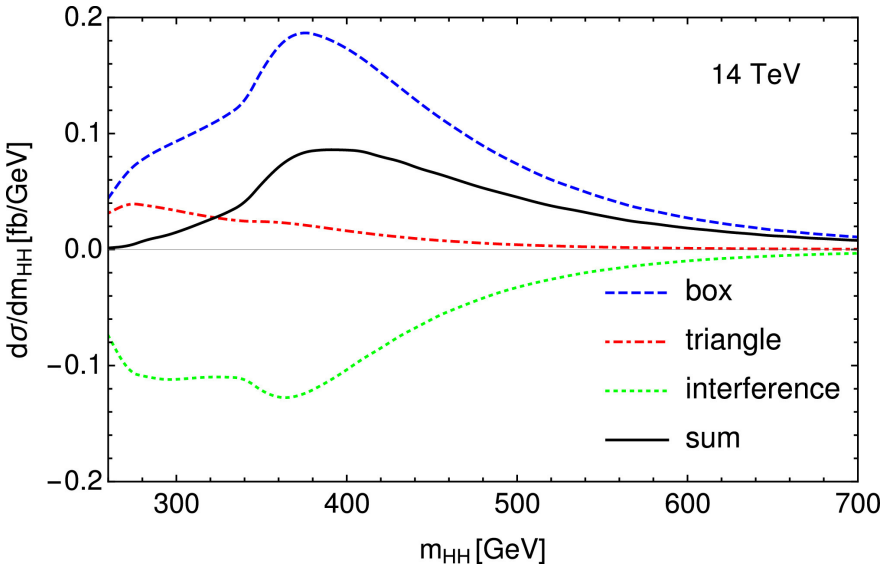


Figure 15.2: m_{HH} distributions for the different components of leading-order ggF di-Higgs production [27].

the distributions at arbitrary values of κ_λ and the distribution at $\kappa_\lambda = 1$. This is done in m_{HH} bins of 10 GeV. The resulting weights are applied to the reconstructed $\kappa_\lambda = 1$ samples for each analysis channel to obtain the shape of the m_{HH} distributions for various κ_λ. The normalization is set using Equation 2.16. $\kappa_\lambda = 10$ samples are generated for each channel to validate the procedure. Validation plots for the $HH \to b\bar{b}b\bar{b}$ channel are shown in Figure 15.4.

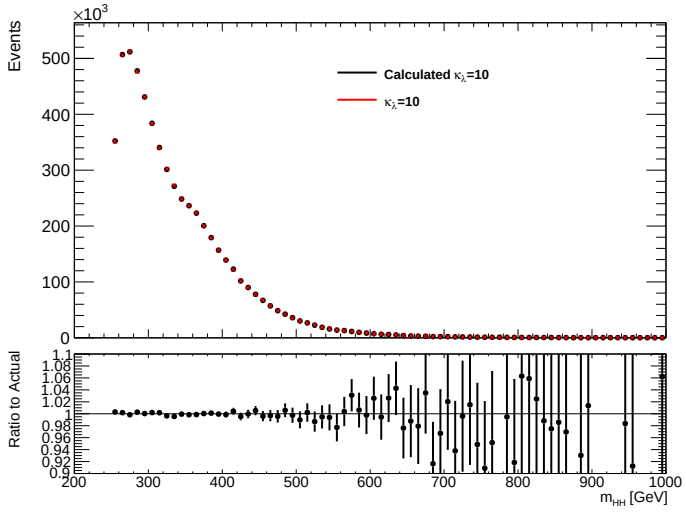
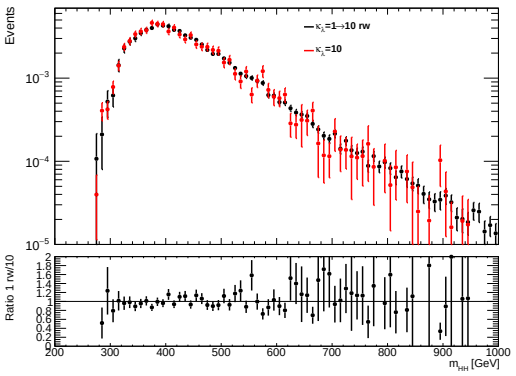
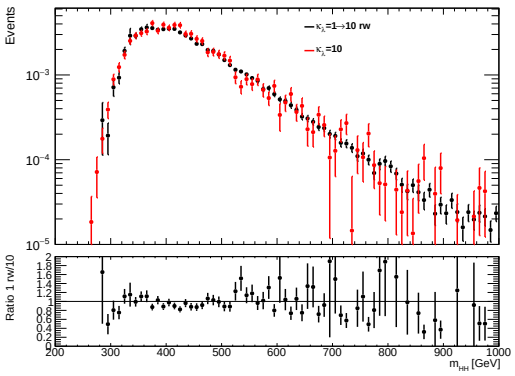


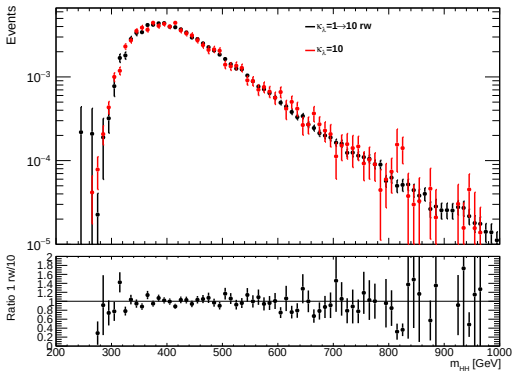
Figure 15.3: Simulated vs. calculated m_{HH} distributions for κ_λ=10.



(a) 2016



(b) 2017



(c) 2018

Figure 15.4: Reweighted vs. simulated $\kappa_\lambda = 10$ m_{HH} distributions for the $HH \to b\bar{b}b\bar{b}$ analysis. Results are generated using Pythia samples with the full analysis selections.

Chapter 16

$HH \to b\bar{b}b\bar{b}$

The $HH \to b\bar{b}b\bar{b}$ channel [104] provides the most statistics of the di-Higgs analyses due to the high $H \to b\bar{b}$ branching ratio. It targets both the ggF and VBF HH production modes. This gives a final state very similar to the SUSY analysis described in Part III, except with no gravitinos to supply missing energy. The analysis thus has very similar backgrounds, with QCD multijet as the dominant background and $t\bar{t}$ as the sub-leading background. We therefore use similar strategies for the two analyses and occasionally share common tools. The below gives a brief overview of the $HH \to b\bar{b}b\bar{b}$ analysis, focusing on the areas where it differs from $HH \to b\bar{b}b\bar{b} + E_{\mathrm{T}}^{\mathrm{miss}}$.

16.1 Object Definitions

As in the SUSY analysis, small-radius particle flow jets with R=0.4 are reconstructed using the anti-k_t algorithm. For the ggF channel, central jets are defined using the same conditions: $p_{\mathrm{T}} \geq 40$ GeV and $|\eta| < 2.5$. However, the VBF channel also uses forward jets, which are defined using $p_{\mathrm{T}} \geq 30$ GeV and $|\eta| \geq 2.5$.

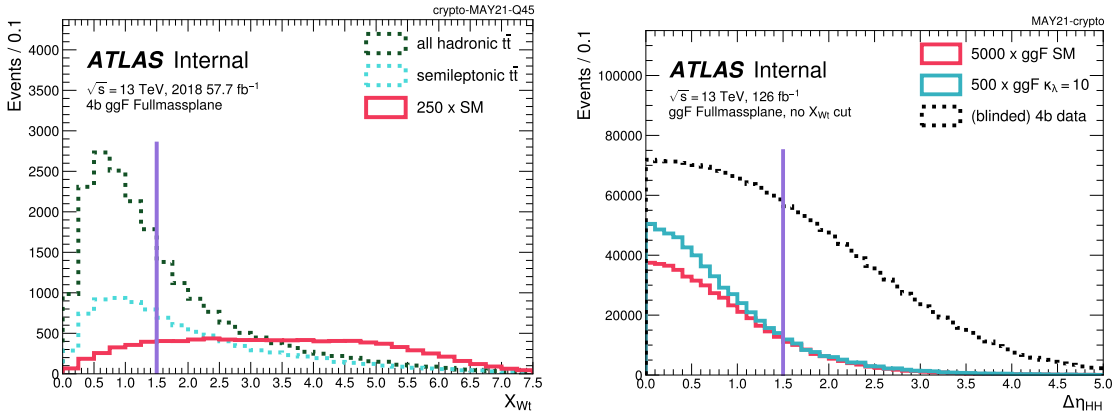
Jet cleaning is applied to events using a tighter JVT working point than for the SUSY analysis, requiring JVT>0.5. Jets are b-tagged using the DL1r algorithm at the 77% working point. Muons are used to correct b-jets using a more sophisticated algorithm than the SUSY analysis, where the closest muon is used if it's within $\Delta R(j, \mu) < \min(0.4, 0.04 + 10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu})$ instead of $\Delta R(j, \mu) = 0.4$. This requires smaller distances for higher energy muons. In addition, correction factors are derived in $t\bar{t}$ to account for energy missed due to b-quarks decaying to neutrinos.

16.2 Event Selection

Higgs boson candidates are reconstructed using the 4 highest p_{T} b-jets in events. If there are less than 4, the remaining jets are selected from the highest p_{T} central non-b-tagged jets in the event (whereas the SUSY analysis uses random central jets). These jets are paired so as the minimize the angular separation between the decay products of the leading Higgs candidate, $\Delta R_{jj,\mathrm{lead}}$.

Events are separated into categories targeting ggF and VBF production. VBF events have ≥ 6 central or forward jets, $m_{jj} > 1$ TeV, and $\Delta\eta_{jj} > 3$ for a pair of non-b-tagged jets, and the p_{T} of the combined 6-jet system must be less than 65 GeV.

The background reduction cuts differ slightly from the SUSY analysis. No lepton veto is applied, and the X_{wt} veto uses a slightly lower threshold of 1.5 instead of 1.8. There are also slight differences in the X_{wt} definition. In this analysis, it only uses jets with $|\eta| < 2.5$, and allows both of the W candidate jets to be Higgs candidate jets. In addition, the ggF channel requires $|\Delta\eta_{HH}| < 1.5$ as the two Higgs bosons are produced from a single interaction point. The efficiency of the X_{wt} and $|\Delta\eta_{HH}|$ cuts for the ggF channel are shown in Figure 16.1.



(a) X_{wt} (b) $|\Delta\eta_{HH}|$

Figure 16.1: X_{wt} and $|\Delta\eta_{HH}|$ distributions for 2018 4b ggF data and MC in the $HH \to b\bar{b}b\bar{b}$ analysis. Events with X_{wt} below the line at 1.5 or $|\Delta\eta_{HH}|$ above the line at 1.5 are removed.

The signal region matches the SUSY analysis except with a slightly offset center,

$$X_{hh}^{\mathrm{SR}} = \sqrt{\left(\frac{m_{h,\mathrm{lead}} - 124\ \mathrm{GeV}}{0.1 m_{h,\mathrm{lead}}}\right)^2 + \left(\frac{m_{h,\mathrm{subl}} - 117\ \mathrm{GeV}}{0.1 m_{h,\mathrm{subl}}}\right)^2} < 1.6 \tag{16.1}$$

Likewise, this center is used for the control regions, where the overall size is also slightly smaller (45 GeV instead of 55 GeV).

$$X_{hh}^{\mathrm{CR}} = \sqrt{(m_{h,\mathrm{lead}} - 124 * 1.05\ \mathrm{GeV})^2 + (m_{h,\mathrm{subl}} - 117 * 1.05\ \mathrm{GeV})^2} < 45\ \mathrm{GeV} \tag{16.2}$$

Instead of having a validation region between the CR and SR, the CR is divided into four quadrants, divided by lines offset 45° from the m_{h1}/m_{h2} axes. The top and bottom quadrants are labeled CR1 and function as the control region while the left and right are labeled CR2 and function as the validation region (VR). The distributions in the Higgs mass plane are shown in Figure 16.2 for signal and Figure 16.3 for data.

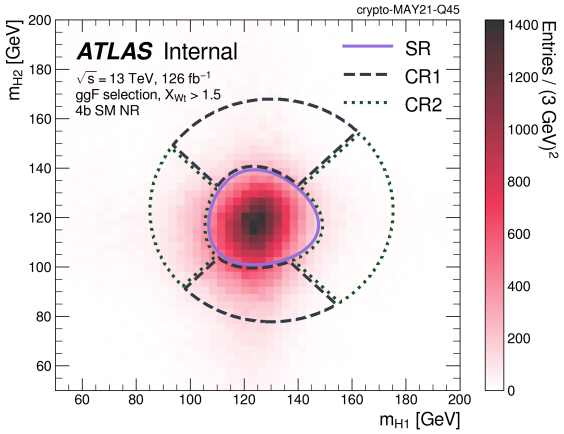
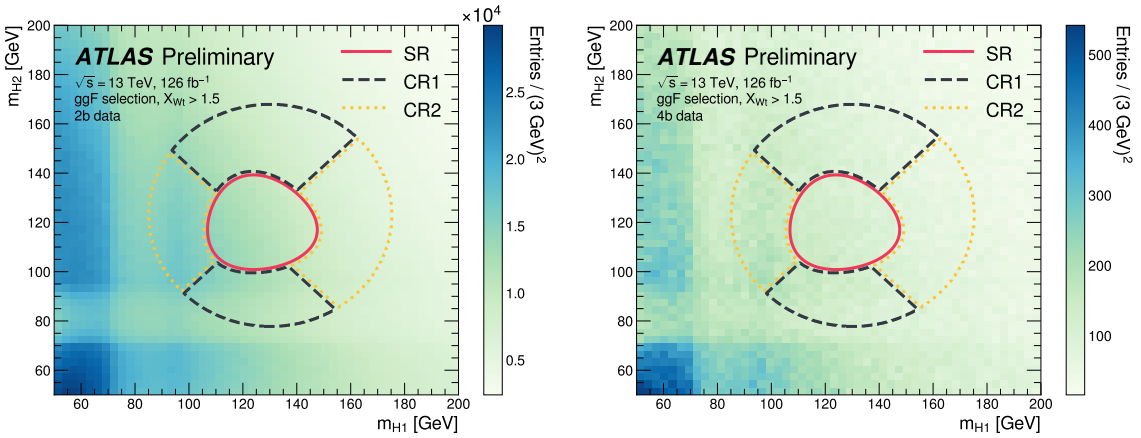


Figure 16.2: Higgs candidate mass distribution for the Standard Model $HH \to b\bar{b}b\bar{b}$ signal using 2016-2018 simulations.

16.3 Trigger Modeling

This analysis uses a subset of the 2b triggers used by the SUSY analysis (see Chapters 7 and 8). Only the 2b2j and ISR triggers are used, while the H_{T} trigger is excluded. A bucket



(a) 2b (b) 4b

Figure 16.3: Higgs candidate mass distributions for the 2b and 4b $HH \to b\bar{b}b\bar{b}$ regions using 2016-2018 data [104].

strategy is employed to create orthogonal trigger categories. Unlike the SUSY analysis, the same thresholds are used for each year: events are placed in the ISR bucket if the leading jet p_{T} is greater than 170 GeV and the p_{T} of the third-leading jet is greater than 70 GeV. All other events are placed in the 2b2j bucket. The trigger efficiencies are shown in Figure 16.4.

The trigger scale factors and their uncertainties are derived in the same way as for the SUSY analysis (Chapter 8).

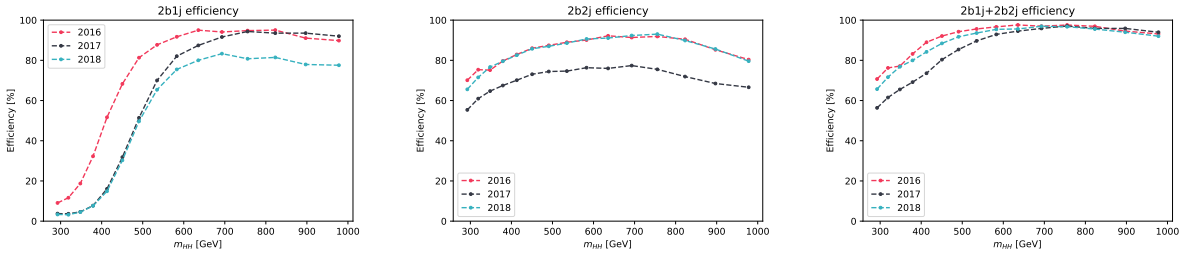


Figure 16.4: Trigger efficiencies as a function of the di-Higgs invariant mass m_{HH}. These are for the Standard Model (κ_λ=1) ggF signal.

16.4 Background Modeling

The background is modeled using the ABCD method. While the SUSY analysis uses boosted decision trees, the di-Higgs analysis uses a neural network (NN). The VBF channel follows a slightly different strategy, reweighting all 3 years inclusively to increase the statistics and applying the top veto before reweighting. 100 variations were generated using the bootstrapping procedure, where the mean of the variations was used. Figure 16.5 shows the reweighting performance for ggF in CR1 while Figure 16.6 shows the reweighting performance for CR2.

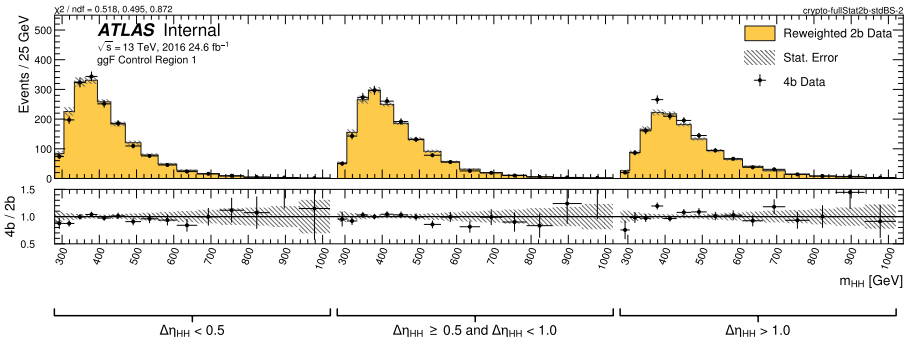
16.5 Categorization and Binning

The ggF channel is divided into 6 categories. There are 3 categories in $|\Delta\eta_{HH}|$: $0 < |\Delta\eta_{HH}| < 0.5$, $0.5 < |\Delta\eta_{HH}| < 1.0$, and $1.0 < |\Delta\eta_{HH}| < 1.5$. Each of these are split into 2 X_{hh}^{SR} categories: $X_{hh}^{\mathrm{SR}} < 0.95$ and $X_{hh}^{\mathrm{SR}} > 0.95$. VBF uses only 2 analysis categories, central ($|\Delta\eta_{HH}| < 1.5$) and forward ($|\Delta\eta_{HH}| > 1.5$). Within each channel and category, a variable-width m_{HH} binning is used to perform the fit.

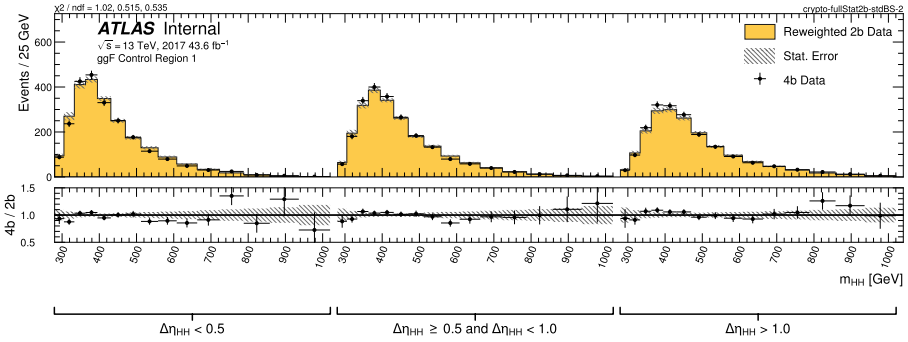
16.6 Systematic Uncertainties

To account for the statistical uncertainty on the reweighting, the standard deviation of the bootstrap variations is set as a bin-by-bin statistical uncertainty. This contrasts with the SUSY analysis, which uses half of the difference between the 84th and 16th percentiles. These are equivalent in the case where the variations are Gaussian.

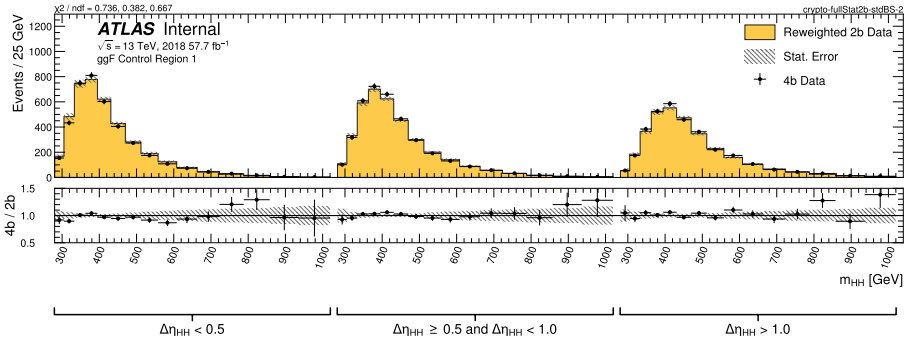
Unlike the SUSY analysis, no non-closure uncertainty is assigned to account for the difference between the 4b and reweighted 2b data in the control region. However, a similar method is used to assign the transfer shape uncertainty. Besides the nominal background estimate (weights derived in CR1), four alternative estimates are made. Like the control



(a) 2016

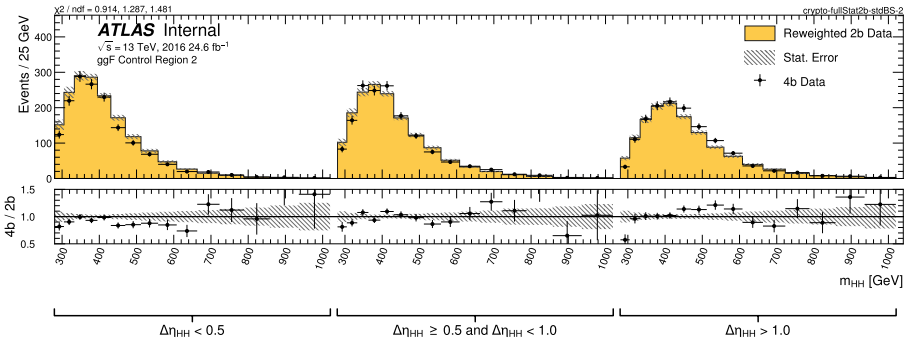


(b) 2017

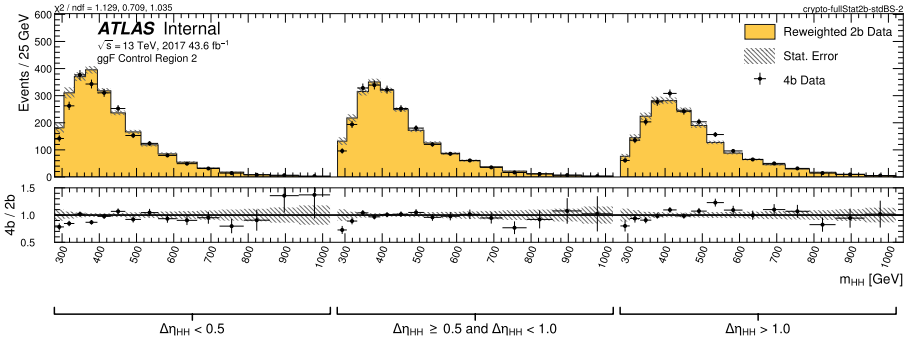


(c) 2018

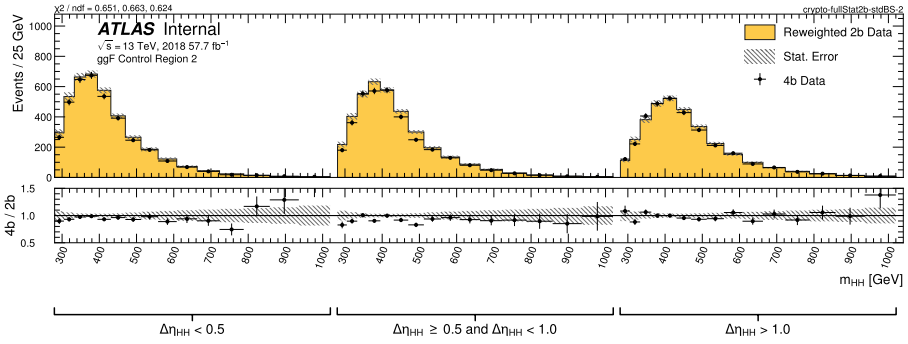
Figure 16.5: 4b vs. reweighted 2b comparison in the 2016, 2017, and 2018 CR1s for the $HH \to b\bar{b}b\bar{b}$ ggF channel.



(a) 2016



(b) 2017



(c) 2018

Figure 16.6: 4b vs. reweighted 2b comparison in the 2016, 2017, and 2018 CR2s for the $HH \to b\bar{b}b\bar{b}$ ggF channel.

region, the signal region is divided into four quadrants (Figure 16.7). For each background estimate, the CR1-derived weights are used in three of the four SR quadrants, while CR2-derived weights are used for the remaining quadrant. The differences between the nominal and alternative estimates are set as background systematics. No separate uncertainty is assigned for normalization, as this is already included in the systematic.

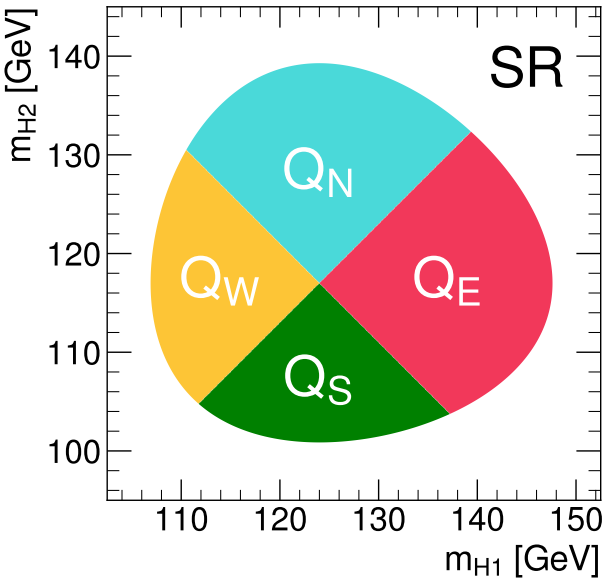


Figure 16.7: Signal region quadrants.

An additional non-closure uncertainty is assigned for the ggF channel. This uses a "3b1f" region with 3 b-jets tagged at the 77% efficiency working point but no other jets tagged at the looser 85% working point. The background estimation procedure is repeated using this region in place of the 4b region. Any differences not covered by the statistical or other systematic uncertainties is set as a non-closure systematic. A summary of the background uncertainties is shown in Figure 16.9.

Detector modeling uncertainties on the jet energy scale, jet energy resolution, flavor tagging, and trigger modeling are included in the same way as in the SUSY analysis. As the background estimate is purely data-driven, these are only applied to signal samples. Systematics on $E_{\mathrm{T}}^{\mathrm{miss}}$ are not used as the analysis does not depend on this variable. Theoretical uncertainties for the $HH \to 4b$ process are applied.

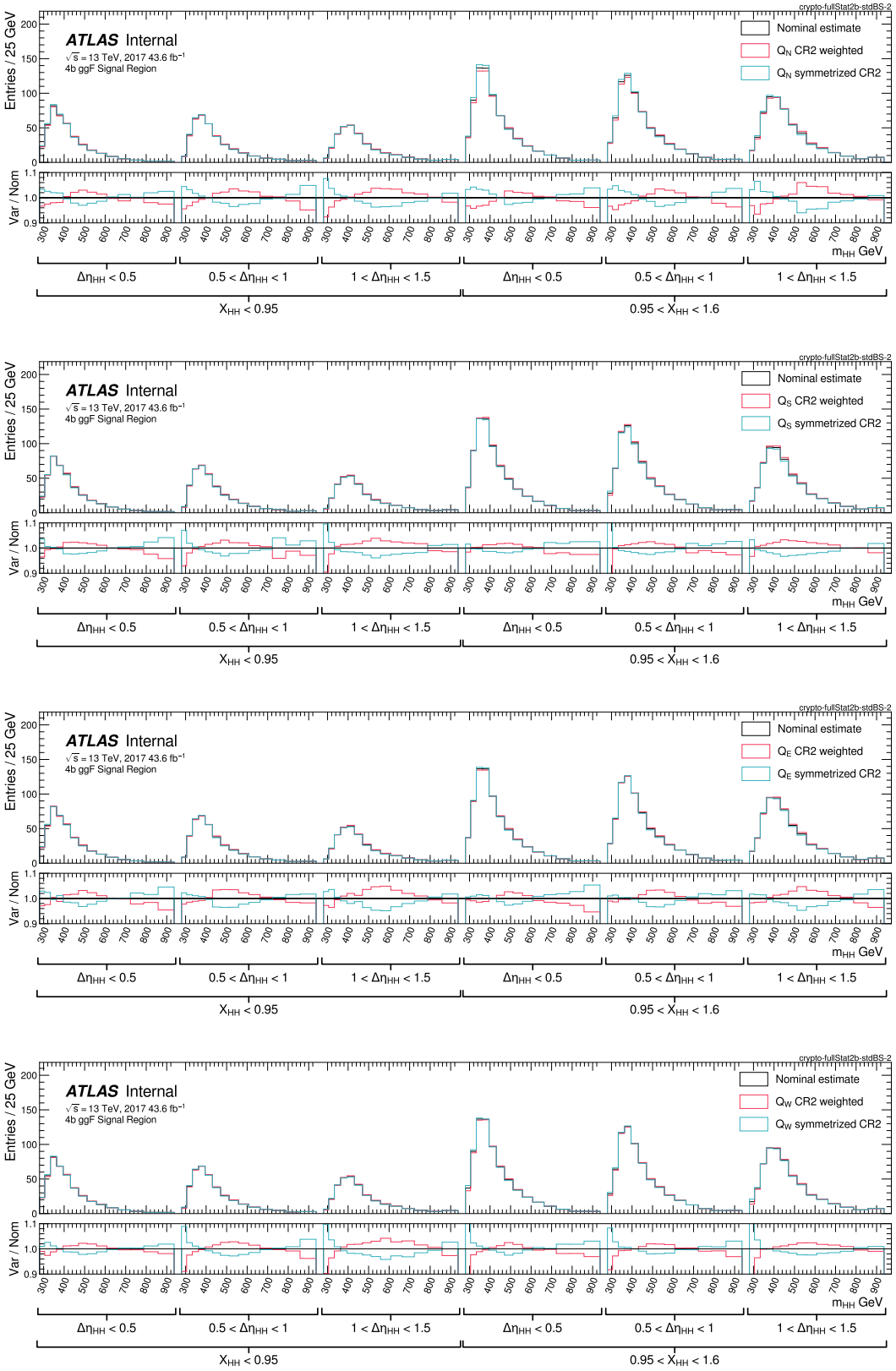
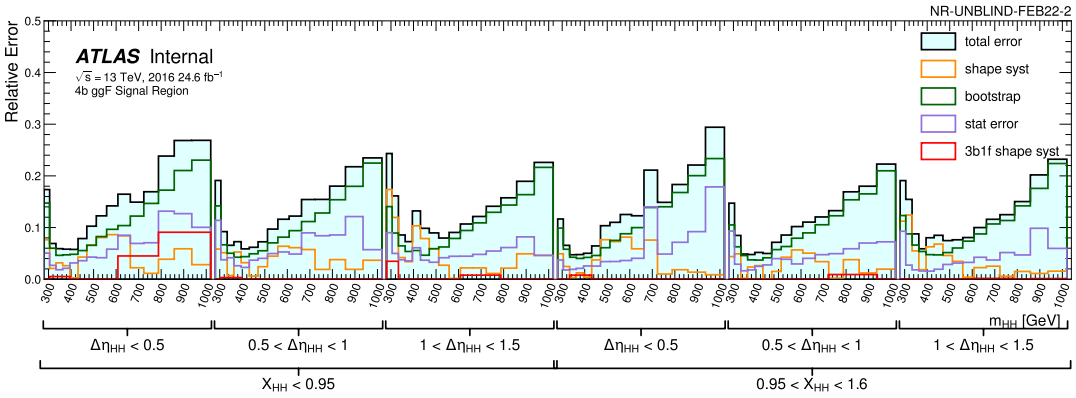
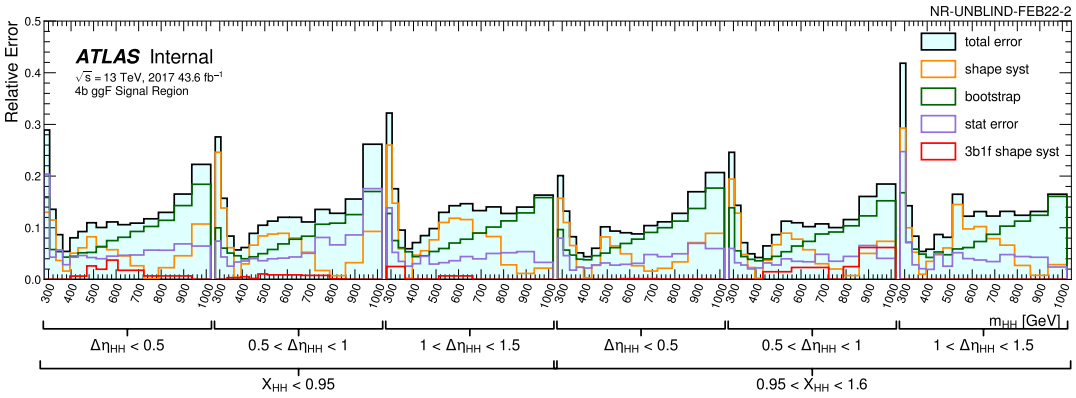


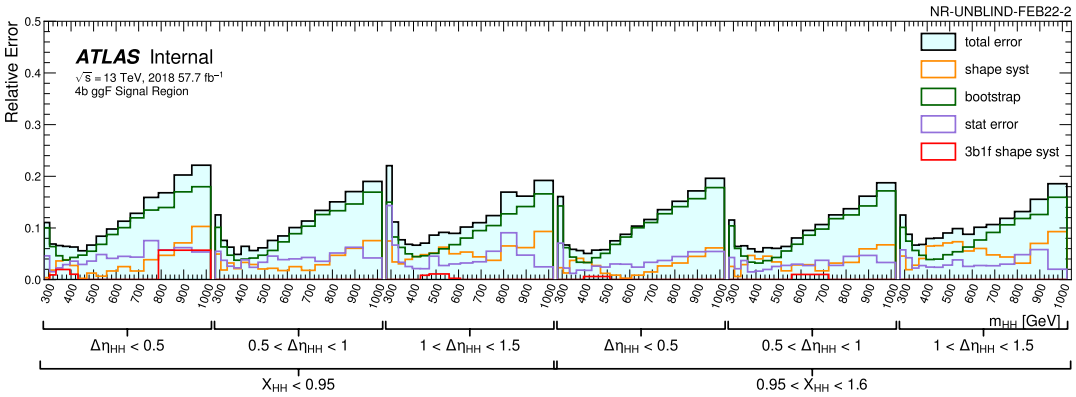
Figure 16.8: Transfer shape uncertainties for the 2017 ggF data in each of the four quadrants.



(a) 2016



(b) 2017



(c) 2018

Figure 16.9: Background uncertainties as a fraction of background yields for the 2016, 2017, and 2018 SRs of the ggF $HH \to b\bar{b}b\bar{b}$ analysis.

16.7 Results

Figure 16.10 shows the 95% confidence level limits as a function of κ_λ. Using the ggF and VBF channels, κ_λ is constrained to fall between -3.9 and 11.1 (-4.6 and 10.8 expected). Alternatively, assuming the Standard Model kinematics, the cross section for Higgs boson pair production is constrained to be less than 5.4 times the Standard Model (8.1 expected).

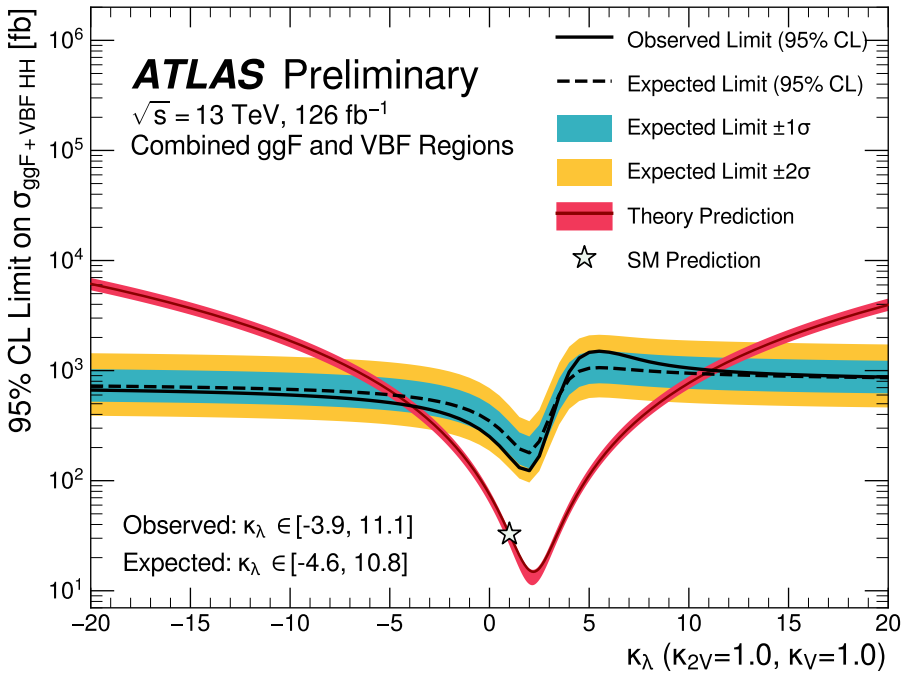


Figure 16.10: 95% CL constraints on κ_λ for the $HH \to b\bar{b}b\bar{b}$ channel [104].

Chapter 17

COMBINATION

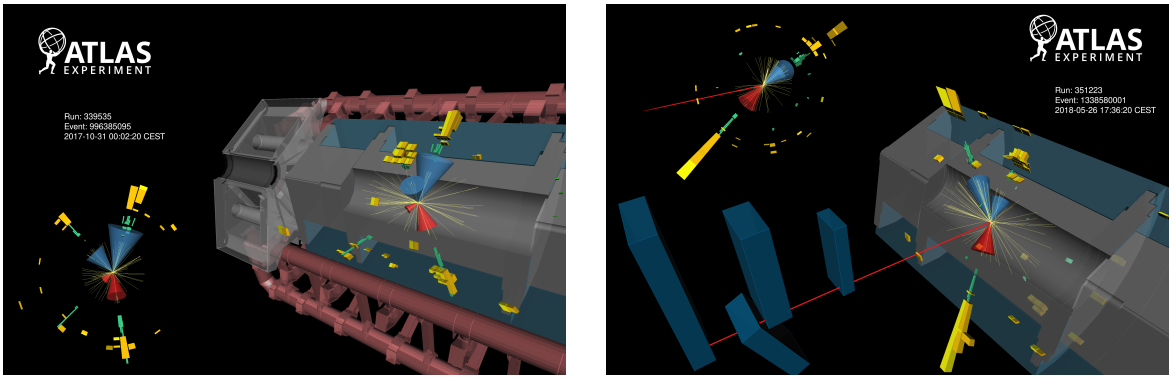
Di-Higgs production is also analyzed in the $HH \to b\bar{b}\gamma\gamma$ and $HH \to b\bar{b}\tau^+\tau^-$ decay modes. Analyses using these channels are briefly summarized below. Along with $HH \to b\bar{b}b\bar{b}$, this gives us a total of 3 independent constraints on Higgs boson pair production. In addition, the results are combined to set the tightest constraints possible. All channels use the κ_λ reweighting described in Chapter 15.

17.1 $HH \to b\bar{b}\tau^+\tau^-$

As shown in Figure 2.11, the $b\bar{b}\tau^+\tau^-$ channel has a relatively high branching ratio of 7.3%. Though this is nearly 80% lower than the 4b branching ratio, the channel can still set competitive limits due to the significantly lower background.

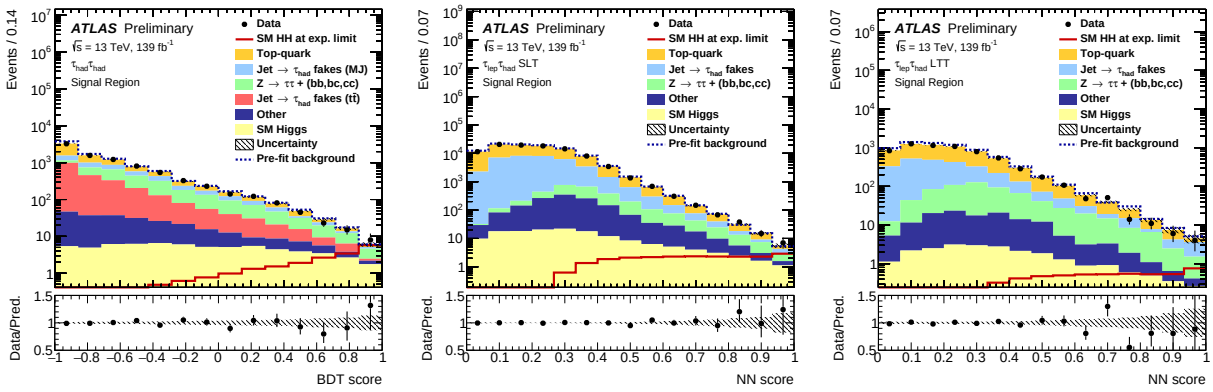
This analysis targets events with 2 b-jets and 2 τ-leptons, where at least one τ must decay hadronically [105]. Events are recorded using a combination of triggers targeting single hadronic τ decays, pairs of hadronic τ decays, single-lepton triggers, and lepton-plus-$\tau_{\text{had-vis}}$ triggers. $\tau_{\text{had-vis}}$ refers to the visible portion of hadronic τ decays. All-hadronic events are placed into one category, while two categories are defined for semi-leptonic events (one for single-lepton trigger (SLT) events and one for lepton-plus-$\tau_{\text{had-vis}}$ trigger (LTT) events). Event displays for the all-hadronic and semi-leptonic channels are shown in Figure 17.1. Hadronic τ candidates are reconstructed using a combination of boosted decision trees and recurrent neural networks (RNNs) [106–108]. Backgrounds with real hadronic τs are estimated using Monte Carlo, while backgrounds with fake τs are estimated with a combination of Monte Carlo and data-driven techniques. A BDT is used to discriminate between signal and background for the all-hadronic category while neural networks are used for the two semi-leptonic categories. The output scores of these multivariate discriminants are shown in Figure 17.2.

This channel limits κ_λ to fall between -2.4 to 9.2 (-2.0 to 9.0 expected) [109]. This is shown in Figure 17.3. An upper limit on the di-Higgs production cross section is set at 4.7 times the Standard Model (3.9 expected) [105].



(a) All-hadronic (b) Semi-leptonic

Figure 17.1: Two candidate events for the $HH \to b\bar{b}\tau^+\tau^-$ analysis. Red cones represent hadronic τ candidates, red lines represent muon candidates, and blue cones represent b-jet candidates [105].



(a) All-hadronic (b) Semi-leptonic, SLT (c) Semi-leptonic, LTT

Figure 17.2: Mutlivariate discriminant score distributions for the $HH \to b\bar{b}\tau^+\tau^-$ analysis. SLT stands for single-lepton trigger and LTT stands for lepton-plus-$\tau_{\text{had-vis}}$ triggers [105].

17.2 $HH \to b\bar{b}\gamma\gamma$

The $b\bar{b}\gamma\gamma$ channel has an even lower branching ratio of 0.26%, over 100 times smaller than the 4b channel. Photons have a very low coupling to the Higgs boson, but yield an

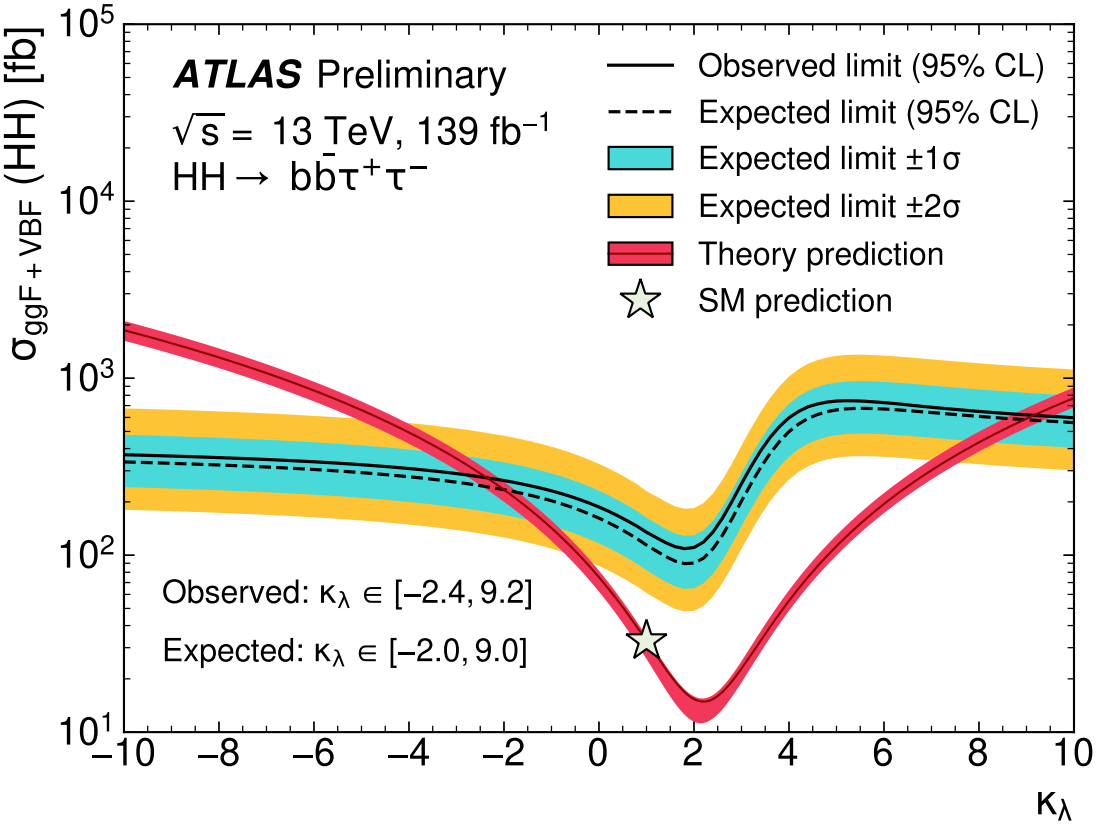


Figure 17.3: 95% CL constraints on κ_λ for the $HH \to b\bar{b}\tau^+\tau^-$ channel [109].

experimentally clean signature. A display of a candidate signal event is shown in Figure 17.4.

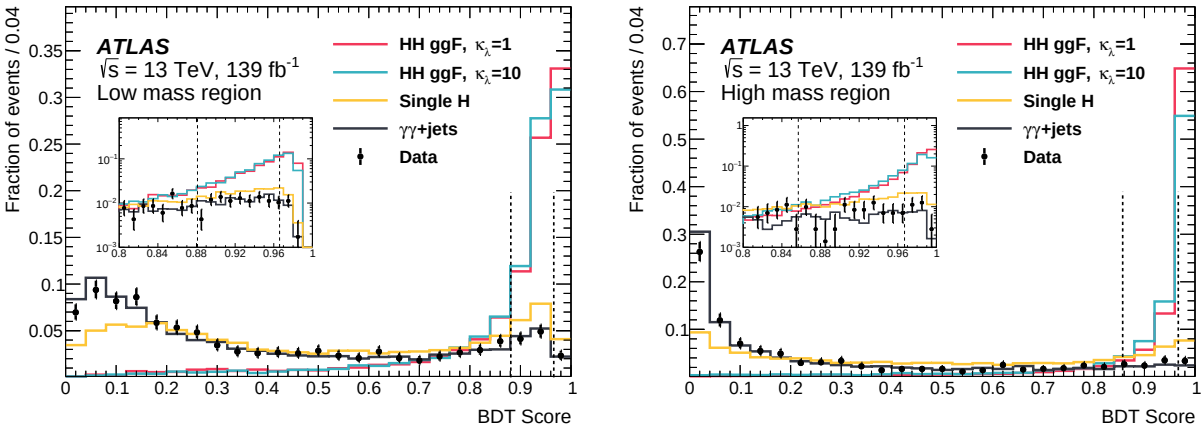
This channel uses di-photon triggers to capture events. The $\gamma\gamma$ background is estimated by fitting a Monte Carlo template to a functional form, and Boosted Decision Trees (Figure 17.5) are used to discriminate between signal and background events. Events are categorized based on their BDT score and reconstructed di-Higgs invariant mass. Results are obtained by fitting on $m_{\gamma\gamma}$, the mass of the Higgs candidate decaying to photons. This benefits from the excellent photon energy resolution of the ATLAS detector [110].

From this analysis, we exclude models outside of $-1.5 < \kappa_\lambda < 6.7$ ($-2.4 < \kappa_\lambda < 7.7$ expected) (see Figure 17.6). Alternatively, the cross section of Higgs boson pair production is constrained to < 4.1 times the Standard Model (5.5 expected) [110].

Figure 17.4: Two views of an event for the $HH \to b\bar{b}\gamma\gamma$ analysis. Tracks are shown in green, b-jet candidates in red, and photons in cyan [110].

17.3 Combination

By combining these analyses, we can place stronger constraints on Higgs boson pair production. The combination includes the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$, and $b\bar{b}b\bar{b}$ channels for di-Higgs production. With these, an upper limit on the cross section σ_{HH} is set at 2.4 times the Standard Model prediction (2.9 expected) (see Figure 17.7) [111]. In addition, κ_λ is restricted to fall between $-0.6 < \kappa_\lambda < 6.6$ ($-2.1 < \kappa_\lambda < 7.8$ expected) (see Figure 17.8). Including



(a) Low-mass region

(b) High-mass Region

Figure 17.5: BDT score distributions for the $HH \to b\bar{b}\gamma\gamma$ channel [110].

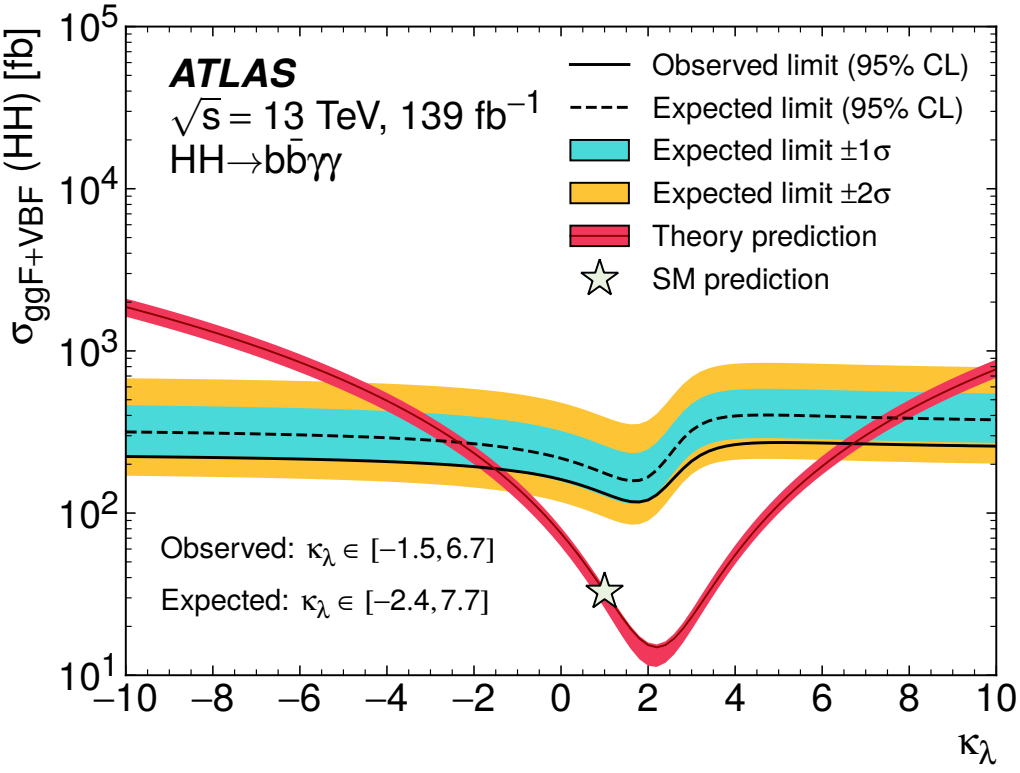


Figure 17.6: 95% CL constraints on κ_λ for the $HH \to b\bar{b}\gamma\gamma$ channel [110].

additional constraints from single Higgs boson production yields slightly stronger limits of $-0.4 < \kappa_\lambda < 6.3$ ($-1.9 < \kappa_\lambda < 7.5$ expected). These are the strongest constraints on κ_λ to date.

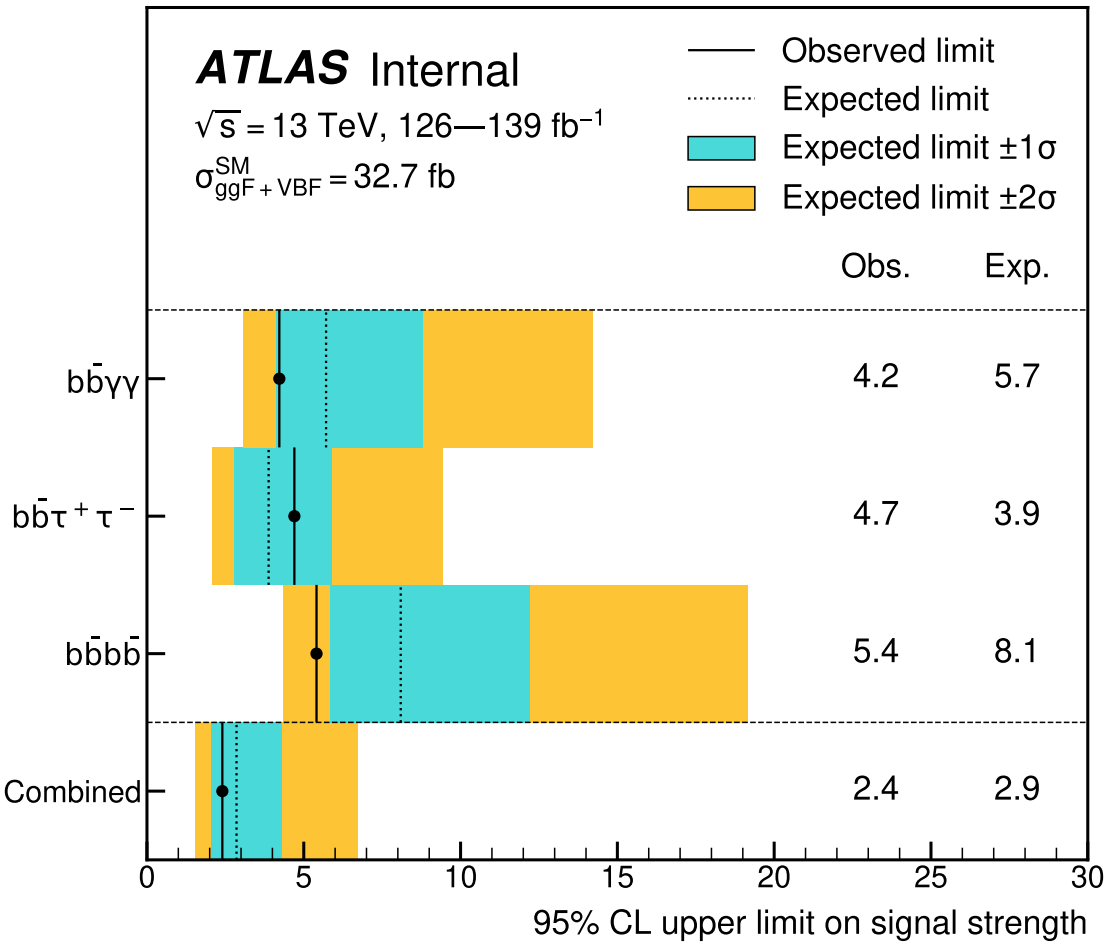


Figure 17.7: 95% CL limits on the cross section of Higgs boson pair production for the various channels and their combination [111].

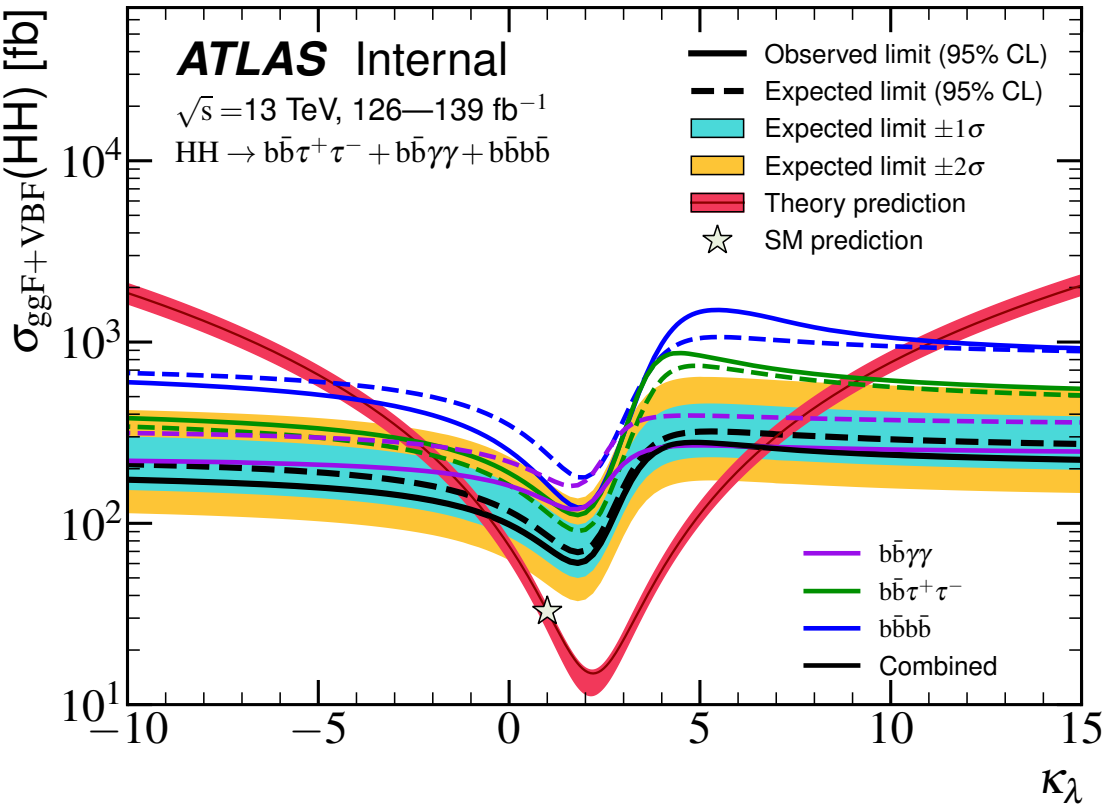


Figure 17.8: 95% CL constraints on κ_λ for the $HH \to b\bar{b}\tau^+\tau^-$, $HH \to b\bar{b}\gamma\gamma$, and $HH \to b\bar{b}b\bar{b}$ channels [111].

Chapter 18

CONCLUSION

Higgs boson pair production is one of the most promising targets for Run 3 and the High-Luminosity LHC (HL-LHC). In a sense, it presents a "no-lose" scenario: if the Standard Model is true, this process exists; if the Standard Model is false, the deviations from the expected self-coupling should be detectable. This is confounded by one unfortunate fact: due to the destructive interference in the gluon-gluon fusion production channel, the SM cross section of 31 fb^{-1} is frustratingly low. This makes the process difficult to detect unless there are large deviations from the Standard Model.

This thesis presents a search for di-Higgs production in the $b\bar{b}b\bar{b}$ decay channel, with a brief overview of the combination with the $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ channels. In addition to the increase in luminosity, many improvements were made from the previous iteration. The background reweighting BDT was upgraded to a neural network, a specialized channel was made for VBF events, and the analysis benefited from improvements to ATLAS jet reconstruction. No significant excess was observed. Upper limits of 5.4 and 2.4 times the SM cross section are set for the $b\bar{b}b\bar{b}$ [104] and combined [109] analyses respectively. Adding in shape information, the self-coupling modifier κ_λ is constrained to $\kappa_\lambda \in [-3.9, 11.1]$ ($b\bar{b}b\bar{b}$ [104]) and $\kappa_\lambda \in [-0.4, 6.3]$ (combination [111]) at 95% CL. While the Standard Model value has not yet been ruled out, these are the most stringent limits to date, reflecting the vast amount of work that has gone into these analyses.

Beyond the increase in luminosity for Run 3 and beyond, there is plenty of room for improvement. Many ideas are in development that did not make this round but could be put to use in future analyses. Some ideas work for both the SUSY analysis (Chapter 14) and $b\bar{b}b\bar{b}$, such as asymmetric b-jet triggers or hardware tracking for the trigger. Even stronger gains are expected for $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$, as these channels are more limited by statistics.

Current projections suggest the HL-LHC could achieve 4σ evidence of di-Higgs production [112]. However, significant work is being done to further optimize each analysis and push

down the systematic uncertainties. With a lot of effort (and a bit of luck) it may be possible to make a 5σ discovery.

Part V

Conclusions

Chapter 19

CONCLUSION

The Higgs boson, and in particular di-Higgs final states, provides a useful probe of the Standard Model. This thesis presented two searches utilizing this process: a search for pair-produced higgsinos and a search for di-Higgs production via the Higgs self-coupling. Despite the vastly different models targeted by these analyses, their similar final states allow analogous strategies to be used. In particular, many advances in either analysis (such as in triggering, Higgs reconstruction, or background estimation) could be applied to the other as well. This motivates continued efforts to work and share techniques between the searches.

As discussed in their respective sections, neither of these analyses yielded evidence of new physics. Despite this, there is reason to be hopeful. While the data taken by ATLAS will only double over Run 3, the High-Luminosity LHC will provide a factor of 10 increase in the integrated luminosity. In the longer term, many future colliders have been proposed that would provide powerful increases in our amount of data and ability to do physics.

However, this does not mean we must sit idly by and wait for new machines. The Run 3 dataset will provide an excellent opportunity to rethink and re-optimize our analyses, as well as to create new ones. The searches presented here have plenty of room for improvement: by implementing more sophisticated tools and strategies, we can move beyond the $\sqrt{2}$ increase in sensitivity expected from the larger dataset. These improvements will be complemented by new techniques to target the current blind spots of our models.

Appendix

Appendix A

NUISANCE PARAMETER BEHAVIOR FOR SUSY SEARCH

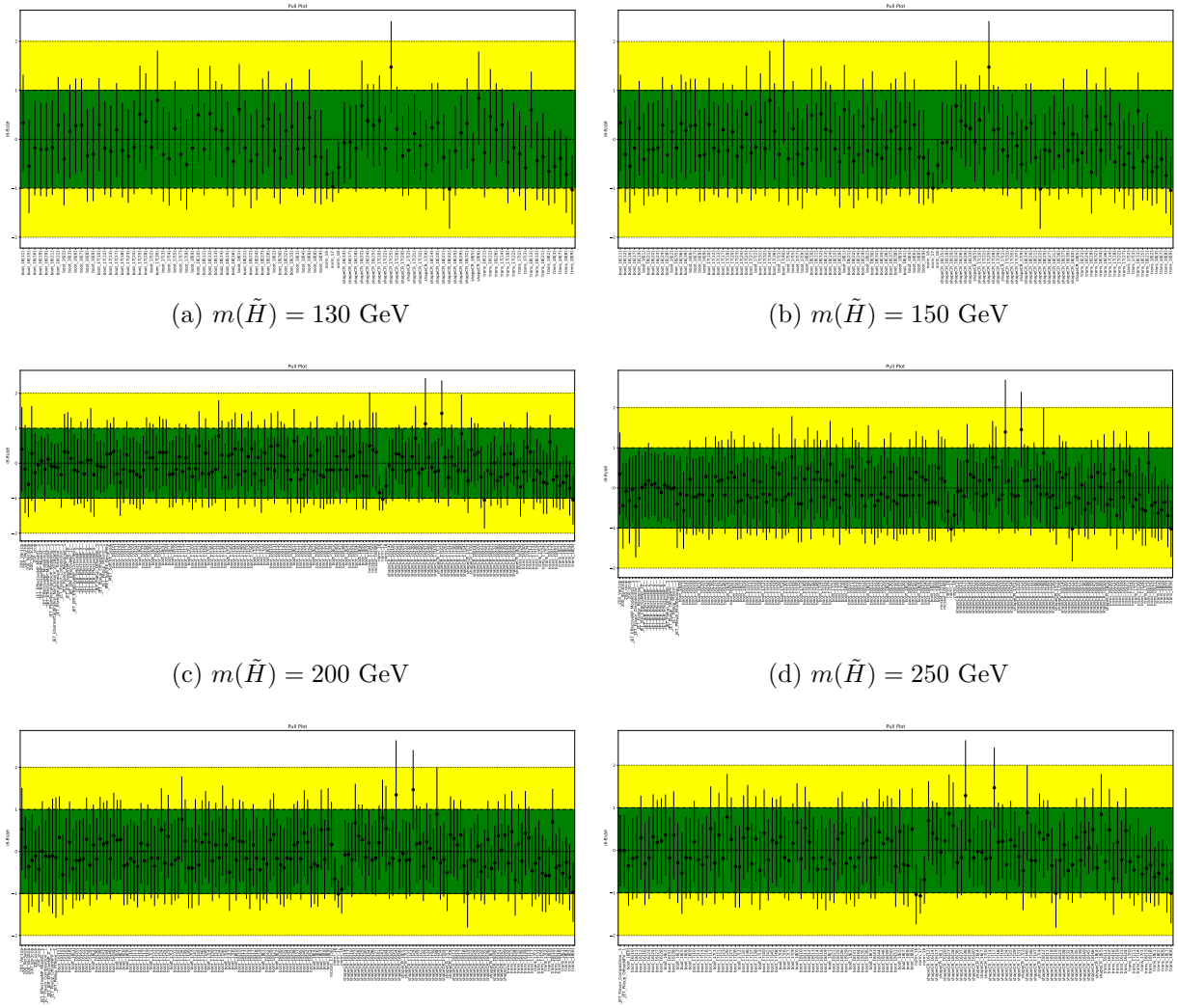
This appendix gives a detailed discussion of the nuisance parameters used in the fit for the SUSY analysis. The nuisance parameters correspond to the systematics defined in Chapter 11. The results of the fit are discussed in Chapter 13.

The non-closure, transfer shape, 2b statistical, and signal MC statistical uncertainties are set as uncorrelated shape systematics. These each give 1 nuisance parameter per bin per channel and use Poisson constraint terms in the statistical model. This analysis has three channels, one for each year of data (2016, 2017, and 2018). The background normalization uncertainties are assigned one NP per channel using Gaussian constraint terms. The detector modeling (including trigger modeling) and signal theory uncertainties are set as correlated shape systematics. These have one NP per systematic, with the variations applied coherently across bins and correlated across channels. Finally, a global luminosity nuisance parameter is assigned with Gaussian constraints.

A.1 Pulls

The pulls for the nuisance parameters are shown in Figure A.1. These show how much the fitted values of the parameters differ from their pre-fit values. The uncertainties are also shown. These can be less than (greater than) 1 due to the fit constraining (anti-constraining) the NPs. Only nuisance parameters with pulls $> 0.15\sigma$, constraints $> 15\%$, or anti-constraints $> 15\%$ are shown in the Figure. Most nuisance parameters are well-behaved, exhibiting typical fluctuations with the 1σ band. The exceptions are the normalization nuisance parameters, which are consistently constrained and fit to $\approx -1\sigma$. These uncertainties were defined by comparing the predicted normalizations with the observed values in the VR. The observed values were lower than the expectations, meaning that a -1σ pull would move the predicted VR normalizations to the observed values. The pulls therefore suggest that

the normalizations in the SR differ from the predictions by similar amounts as in the VR.



(a) $m(\tilde{H}) = 130$ GeV

(b) $m(\tilde{H}) = 150$ GeV

(c) $m(\tilde{H}) = 200$ GeV

(d) $m(\tilde{H}) = 250$ GeV

(e) $m(\tilde{H}) = 300$ GeV

(f) $m(\tilde{H}) = 500$ GeV

Figure A.1: Pulls for fits using the listed signal masses. Only nuisance parameters constrained more than 15%, anti-constrained more than 15%, or pulled more than 0.15σ are shown. The bracketed values are bin numbers, with the numbering starting at 0.

A.2 Correlations

The correlations between nuisance parameters are shown in Figures A.2-A.6. Only nuisance parameters with correlations (besides self-correlations) above threshold values are shown. Background and signal systematics are shown in separate plots as any correlations

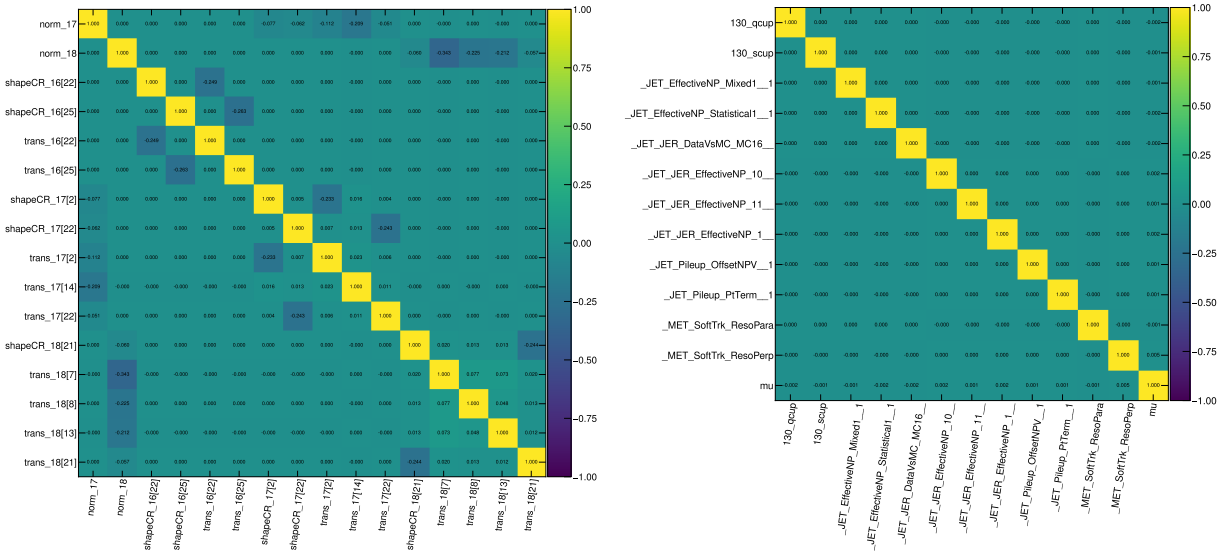
between them are negligible. Large correlations (or anti-correlations) could suggest that our model is insufficient, as it assumes distinct nuisance parameters are uncorrelated.

As expected, no significant correlations are observed between background systematics corresponding to different years of data. Slight anti-correlations are observed between the normalization systematics and some non-closure and transfer shape systematics. This is reasonable as either the normalization or shape systematics can increase the background estimate in a bin, so an increase in one could cancel an increase in the other. The worst such anti-correlation observed for the 130-300 GeV fits is -0.367, occurring between the 2018 normalization systematic and one of the transfer shape systematics for the 300 GeV fit.

Significant correlations and anti-correlations are also observed between the signal strength μ and various detector modeling uncertainties. This makes sense for the same reason as the background normalization and shape systematics: the amount of signal seen in a given bin can be increased by varying either μ or a signal shape uncertainty. The largest such correlation is 0.212 for μ and Var1 signal theory uncertainty in the 300 GeV fit. The largest such anti-correlation is -0.318 for μ and the jet flavor composition in the 250 GeV fit.

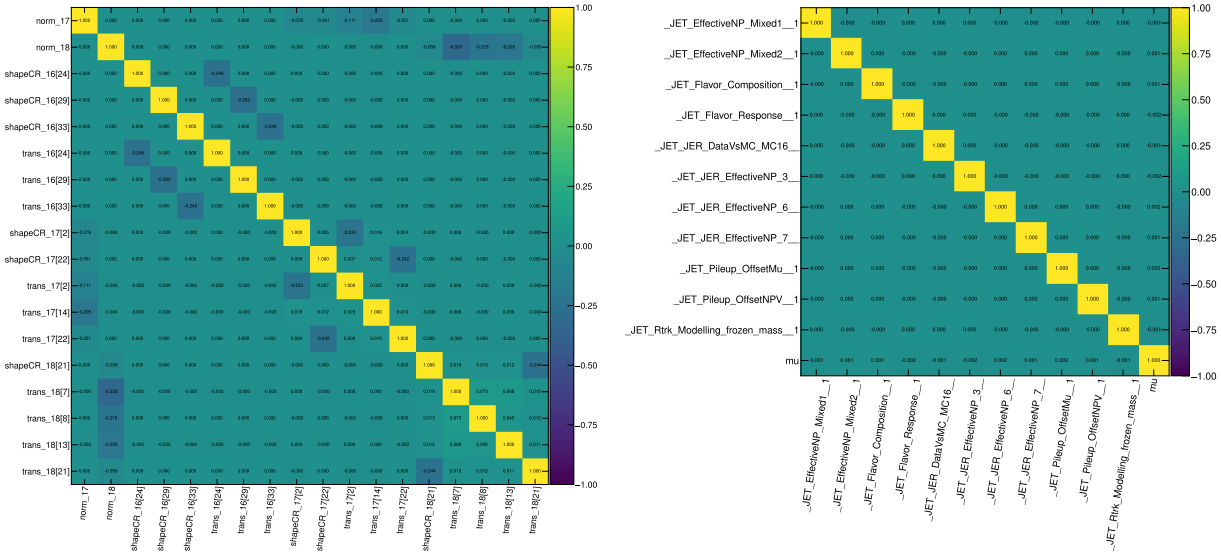
A.3 Impacts

The impact of the nuisance parameters on the results can also be measured. The fit was re-run 4 times for each NP, fixing it at its $\pm 1\sigma$ pre-fit and $\pm 1\sigma$ post-fit values. For each fit, the impact of the nuisance parameter is defined as the difference between the best-fit signal strength μ in the fit and its value in the nominal fit. The results are shown in Figures A.7-A.9. Some of the largest impacts are due to non-closure uncertainties from high $E_{\mathrm{T}}^{\mathrm{miss}}$ bins and signal theory uncertainties. All systematics for the 130 GeV and 150 GeV mass points have negligible impact, except for a couple of pre-fit systematics. The data for these fits are not well described by the signal plus background model regardless of nuisance parameter values, causing the best-fit signal strength μ to be close to 0 in nearly all scenarios.



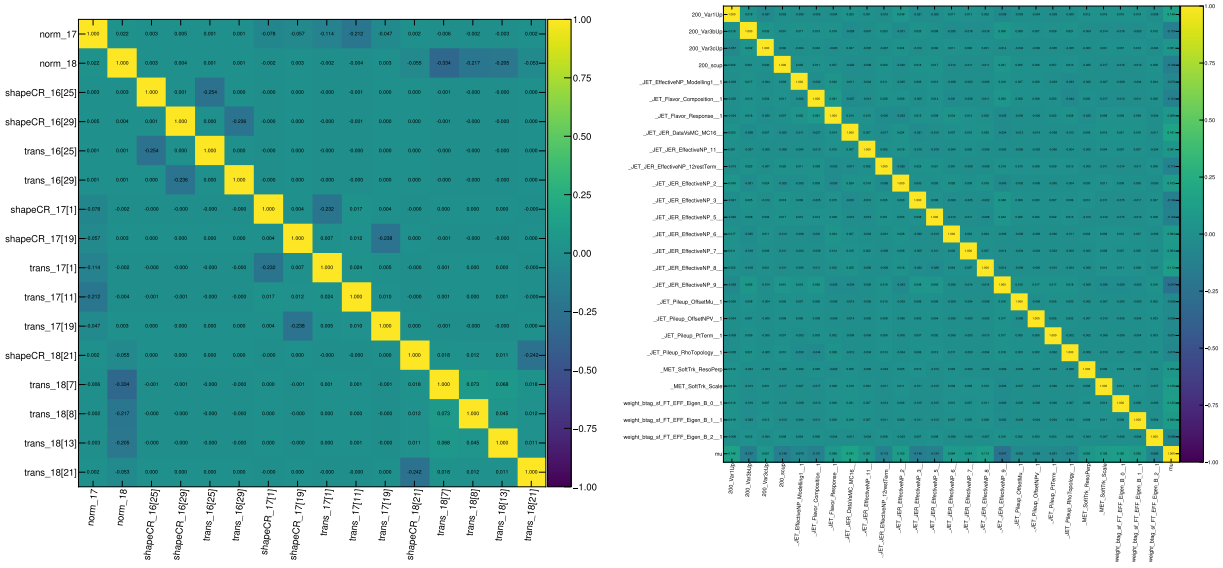
(a) Background sytematics (b) Signal sytematics

Figure A.2: Correlation matrices for fits with free-floating signal strength for the 130 GeV signal. Only nuisance parameters with correlations (besides self-correlation) with magnitude greater than 0.2 (0.001) are shown for the background (signal) systematics.



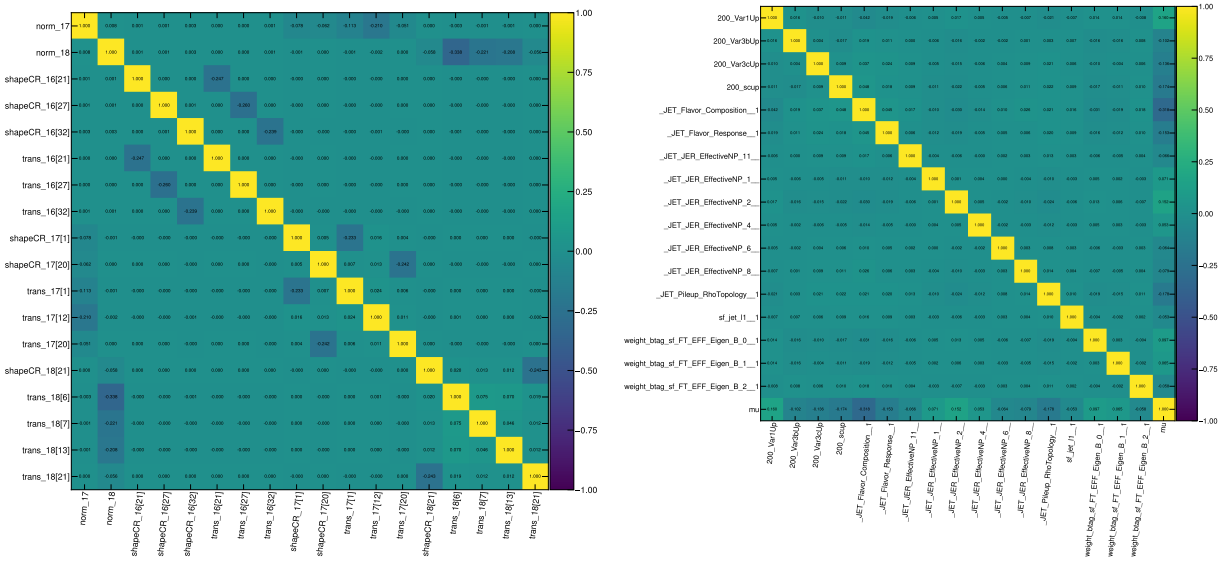
(a) Background sytematics (b) Signal sytematics

Figure A.3: Correlation matrices for fits with free-floating signal strength for the 150 GeV signal. Only nuisance parameters with correlations (besides self-correlation) with magnitude greater than 0.2 (0.001) are shown for the background (signal) systematics.



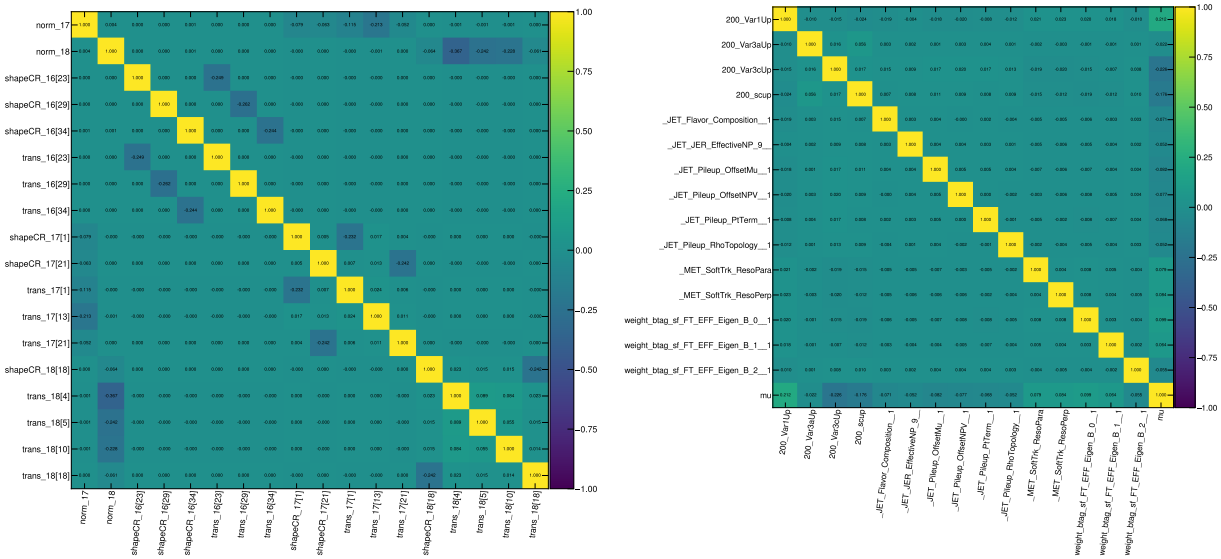
(a) Background sytematics (b) Signal sytematics

Figure A.4: Correlation matrices for fits with free-floating signal strength for the 200 GeV signal. Only nuisance parameters with correlations (besides self-correlation) with magnitude greater than 0.2 (0.05) are shown for the background (signal) systematics.



(a) Background sytematics (b) Signal sytematics

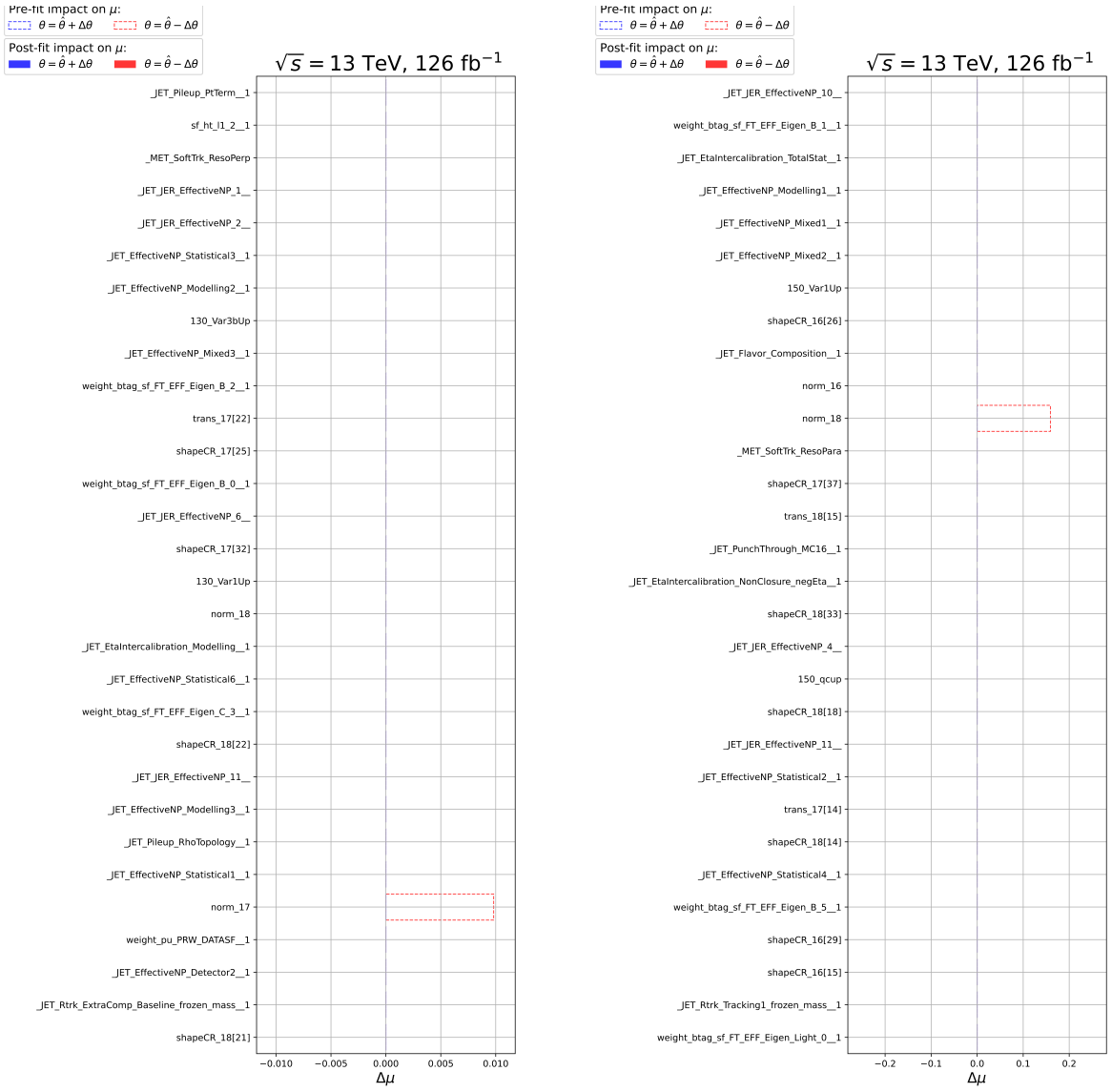
Figure A.5: Correlation matrices for fits with free-floating signal strength for the 250 GeV signal. Only nuisance parameters with correlations (besides self-correlation) with magnitude greater than 0.2 (0.05) are shown for the background (signal) systematics.



(a) Background sytematics

(b) Signal sytematics

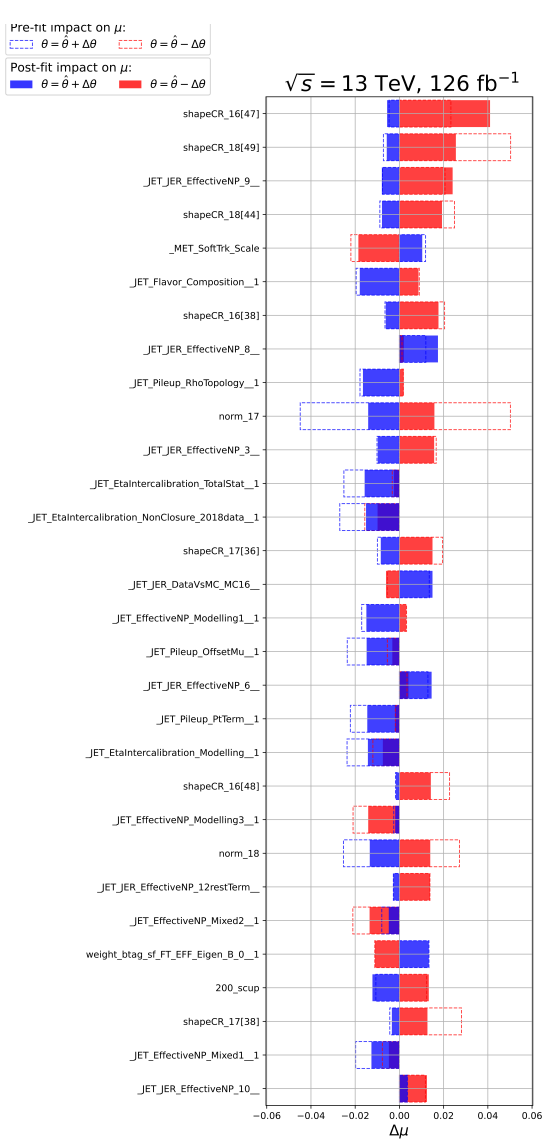
Figure A.6: Correlation matrices for fits with free-floating signal strength for the 300 GeV signal. Only nuisance parameters with correlations (besides self-correlation) with magnitude greater than 0.2 (0.05) are shown for the background (signal) systematics.



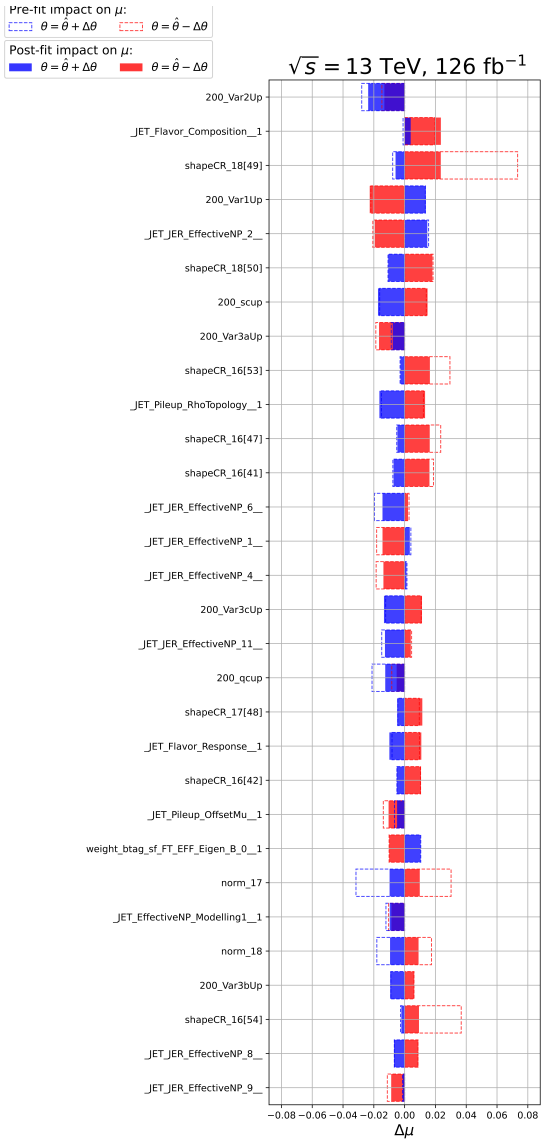
(a) $m(\tilde{H}) = 130$ GeV

(b) $m(\tilde{H}) = 150$ GeV

Figure A.7: Impacts for fits using the 130 GeV and 150 GeV signals. Only the 30 nuisance parameters with the largest maximum post-fit impacts are shown. Pre-fit impacts are dashed while post-fit impacts are solid. Blue corresponds to fixing the value of a parameter at its $+1\sigma$ value and red corresponds to fixing it at its -1σ value. If any of these would result in a multiplicative factor of less than 1e-10 for an uncorrelated shape systematic, 1e-10 is used. The bracketed values are bin numbers, with the numbering starting at 0. These bin numbers only count bins where the nominal signal is non-zero.



(a) $m(\tilde{H}) = 200$ GeV



(b) $m(\tilde{H}) = 250$ GeV

Figure A.8: Impacts for fits using the 200 GeV and 250 GeV signals. Only the 30 nuisance parameters with the largest maximum post-fit impacts are shown. Pre-fit impacts are dashed while post-fit impacts are solid. Blue corresponds to fixing the value of a parameter at its $+1\sigma$ value and red corresponds to fixing it at its -1σ value. If any of these would result in a multiplicative factor of less than 1e-10 for an uncorrelated shape systematic, 1e-10 is used. The bracketed values are bin numbers, with the numbering starting at 0. These bin numbers only count bins where the nominal signal is non-zero.

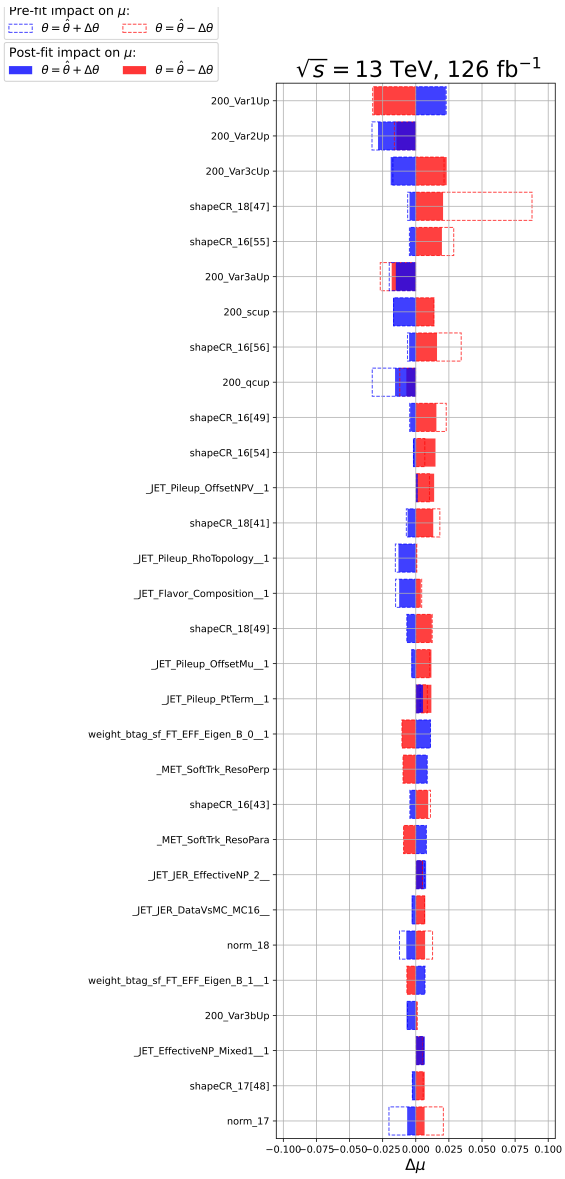


Figure A.9: Impacts for fits using the 300 GeV signal. Only the 30 nuisance parameters with the largest maximum post-fit impacts are shown. Pre-fit impacts are dashed while post-fit impacts are solid. Blue corresponds to fixing the value of a parameter at its $+1\sigma$ value and red corresponds to fixing it at its -1σ value. If any of these would result in a multiplicative factor of less than 1e-10 for an uncorrelated shape systematic, 1e-10 is used. The bracketed values are bin numbers, with the numbering starting at 0. These bin numbers only count bins where the nominal signal is non-zero.

REFERENCES

[1] S. Berryman, "Democritus", *The Stanford Encyclopedia of Philosophy*, ed. by E. N. Zalta, Winter 2016, Metaphysics Research Lab, Stanford University, 2016.

[2] J. J. Thomson, *On the Structure of the Atom: an Investigation of the Stability and Periods of Oscillation of a number of Corpuscles arranged at equal intervals around the Circumference of a Circle; with Application of the Results to the Theory of Atomic Structure*, Philosophical Magazine, 6th ser. **7** (1904) 237.

[3] D. Griffiths, *Introduction to Elementary Particles*, Second, Revised, Wiley-VCH, 2008.

[4] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex].

[5] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex].

[6] S. F. Novaes, "Standard model: An Introduction", *10th Jorge Andre Swieca Summer School: Particle and Fields*, 1999 5, arXiv: hep-ph/0001283.

[7] P. Gagnon, *The Standard Model: a beautiful but flawed theory*, 2014, URL: https://www.quantumdiaries.org/2014/03/14/the-standard-model-a-beautiful-but-flawed-theory.

[8] W. Pauli, *The Connection Between Spin and Statistics*, Phys. Rev. **58** (1940) 716.

[9] Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv: hep-ex/9807003.

[10] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou, *The fate of hints: updated global analysis of three-flavor neutrino oscillations*, JHEP **09** (2020) 178, arXiv: 2007.14792 [hep-ph].

[11] P. A. Zyla et al., *Review of Particle Physics*, PTEP **2020** (2020) 083C01, and 2021 update.

[12] V. A. Rubakov and M. E. Shaposhnikov, *Electroweak baryon number non-conservation in the early universe and in high-energy collisions*, AIP Conf. Proc. **419** (1998) 347, ed. by H. Falomir, R. E. Gamboa Saravi, and F. A. Schaposnik.

[13] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321, ed. by J. C. Taylor.

[14] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[15] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508, ed. by J. C. Taylor.

[16] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585, ed. by J. C. Taylor.

[17] J. Ellis, "Higgs Physics", *2013 European School of High-Energy Physics*, 2015 117, arXiv: 1312.5672 [hep-ph].

[18] S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*, (2013), URL: https://cds.cern.ch/record/1559921.

[19] S. P. Martin, *A Supersymmetry primer*, Adv. Ser. Direct. High Energy Phys. **18** (1998) 1, ed. by G. L. Kane, arXiv: hep-ph/9709356.

[20] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012) 1, arXiv: 1106.0034 [hep-ph].

[21] D. J. H. Chung et al., *The Soft supersymmetry breaking Lagrangian: Theory and applications*, Phys. Rept. **407** (2005) 1, arXiv: hep-ph/0312378.

[22] K. J. Bae, H. Baer, V. Barger, and D. Sengupta, *Revisiting the SUSY μ problem and its solutions in the LHC era*, Phys. Rev. D **99** (2019) 115027, arXiv: 1902.10748 [hep-ph].

[23] P. Meade, N. Seiberg, and D. Shih, *General Gauge Mediation*, Prog. Theor. Phys. Suppl. **177** (2009) 143, ed. by T. Takayanagi and S. Yahikozawa, arXiv: 0801.3278 [hep-ph].

[24] R. Barbieri and D. Pappadopulo, *S-particles at their naturalness limits*, JHEP **10** (2009) 061, arXiv: 0906.4546 [hep-ph].

[25] P. Meade, M. Reece, and D. Shih, *Prompt Decays of General Neutralino NLSPs at the Tevatron*, JHEP **05** (2010) 105, arXiv: 0911.4130 [hep-ph].

[26] ATLAS Collaboration, *Search for pair production of higgsinos in final states with at least three b-tagged jets in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, Phys. Rev. D **98** (2018) 092002, arXiv: 1806.04030 [hep-ex].

[27] J. Alison et al., *Higgs boson potential at colliders: Status and perspectives*, Rev. Phys. **5** (2020) 100045, ed. by B. Di Micco, M. Gouzevitch, J. Mazzitelli, and C. Vernieri, arXiv: 1910.00012 [hep-ph].

[28] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[29] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, *On the Anomalous Electroweak Baryon Number Nonconservation in the Early Universe*, Phys. Lett. B **155** (1985) 36.

[30] D. E. Morrissey and M. J. Ramsey-Musolf, *Electroweak baryogenesis*, New J. Phys. **14** (2012) 125003, arXiv: 1206.2942 [hep-ph].

[31] D. Buttazzo et al., *Investigating the near-criticality of the Higgs boson*, JHEP **12** (2013) 089, arXiv: 1307.3536 [hep-ph].

[32] *LHC Higgs Cross Section HH Sub-group*, URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWGHH?rev=74.

[33] ATLAS Collaboration, *Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **01** (2019) 030, arXiv: 1804.06174 [hep-ex].

[34] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector*, Phys. Lett. B **800** (2020) 135103, arXiv: 1906.02025 [hep-ex].

[35] *LHC Machine*, JINST **3** (2008) S08001, ed. by L. Evans and P. Bryant.

[36] *Experiments*, URL: https://home.cern/science/experiments.

[37] J. Vollaire et al., *Linac4 design report*, ed. by M. Vretenar, vol. 6/2020, CERN Yellow Reports: Monographs, CERN, 2020, ISBN: 978-92-9083-579-0, 978-92-9083-580-6.

[38] *CERN's accelerator complex*, URL: https://home.cern/science/accelerators/accelerator-complex.

[39] H. Jarlett, *25 years of Large Hadron Collider experimental programme*, 2017, URL: https://home.cern/news/news/cern/25-years-large-hadron-collider-experimental-programme.

[40] J. Thomason, R. Garoby, S. Gilardoni, L. Jenner, and J. Pasternak, *Proton driver scenarios at CERN and Rutherford Appleton Laboratory*, Physical Review Special Topics - Accelerators and Beams **16** (2013).

[41] *Facts and figures about the LHC*, URL: https://home.cern/resources/faqs/facts-and-figures-about-lhc.

[42] ATLAS Collaboration, *Configuration and performance of the ATLAS b-jet triggers in Run 2*, Eur. Phys. J. C **81** (2021) 1087, arXiv: 2106.03584 [hep-ex].

[43] A. Das and T. Ferbel, *Introduction to Nuclear and Particle Physics*, Second, World Scientific, 2003.

[44] A. El-Saftawy, "Regulating the performance parameters of accelerated particles", PhD thesis, 2013.

[45] G. J. Wiener, J. Woithe, A. Brown, and K. Jende, *Introducing the LHC in the classroom: an overview of education resources available*, Phys. Educ. **51** (2016) 035001.

[46] X. C. Vidal and R. C. Manzano, *Buckets and Bunches*, URL: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.buckets_and_bunches.

[47] L. Rossi, *The LHC superconducting magnets*, Conf. Proc. C **030512** (2003) 141.

[48] *Pulling together: Superconducting electromagnets*, URL: https://home.cern/science/engineering/pulling-together-superconducting-electromagnets.

[49] J.-L.Caron, "Cross section of LHC dipole.. Dipole LHC: coupe transversale.", AC Collection. Legacy of AC. Pictures from 1992 to 2002., 1998, URL: https://cds.cern.ch/record/841539.

[50] X. C. Vidal and R. C. Manzano, *Magnetic multipoles*, URL: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.magnetic_multipoles.

[51] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[52] A. M. Rodriguez Vera, S. Mehlhase, and K. Anthony, *The ATLAS Collaboration - ATLAS Fact Sheet*, (2021), URL: https://cds.cern.ch/record/2759656.

[53] J. Pequenao and P. Schaffner, "How ATLAS detects particles: diagram of particle paths in the detector", 2013, URL: https://cds.cern.ch/record/1505342.

[54] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, tech. rep., 2010, URL: https://cds.cern.ch/record/1291633.

[55] A. Polini et al., *Design of the ATLAS IBL Readout System*, Phys. Procedia **37** (2012) 1948, ed. by Ted Liu.

[56] G. Knoll, *Radiation Detection and Measurement (4th ed.)* John Wiley, 2010, ISBN: 978-0-470-13148-0.

[57] P. Allport et al., *ATLAS irradiation studies of n-in-n and p-in-n silicon microstrip detectors*, Nucl. Instrum. Meth. A **435** (1999) 74, ed. by D. Loukas and C. Markou.

[58] E. Abat et al., *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, JINST **3** (2008) P02013.

[59] F. Cavallari, *Performance of calorimeters at the LHC*, J. Phys. Conf. Ser. **293** (2011) 012001, ed. by Y. Wang.

[60] S. Palestini, *The muon spectrometer of the ATLAS experiment (talk presented at the "8th Topical Seminar on Innovative Particle and Radiation Detectors", 21-24 Oct. 2002, Siena, Italy)*, tech. rep., Contribution to the proceedings : 8th Topical Seminar on Innovative Particle and Radiation Detectors, 21-24 Oct. 2002, Siena, Italy: CERN, 2003, URL: https://cds.cern.ch/record/681459.

[61] ATLAS Collaboration, *ATLAS muon spectrometer: Technical design report*, (1997).

[62] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[63] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex].

[64] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: 1907.05120 [hep-ex].

[65] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, tech. rep., CERN, 2017, URL: https://cds.cern.ch/record/2273281.

[66] ATLAS Collaboration, *Performance of 2019 recommendations of ATLAS Flavor Tagging algorithms with Variable Radius track jets*, 2019, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2019-006/.

[67] A. M. Rodriguez Vera, S. Mehlhase, R. Gonzalez Suarez, and K. Anthony, *Trigger and Data Acquisition - ATLAS Fact Sheet*, (2021), URL: https://cds.cern.ch/record/2775198.

[68] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[69] R. Achenbach et al., *The ATLAS Level-1 Calorimeter Trigger*, tech. rep., CERN, 2008, URL: https://cds.cern.ch/record/1080560.

[70] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004, arXiv: 2007.12539 [physics.ins-det].

[71] ATLAS Collaboration, *The ATLAS Fast TracKer system*, JINST **16** (2021) P07006, arXiv: 2101.05078 [physics.ins-det].

[72] *Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System*, tech. rep., CERN, 2017, URL: https://cds.cern.ch/record/2285584.

[73] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph].

[74] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012, URL: https://cds.cern.ch/record/1474107.

[75] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV*, JHEP **10** (2012) 081, arXiv: 1207.2159 [hep-ph].

[76] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Precision predictions for electroweak superpartner production at hadron colliders with Resummino*, Eur. Phys. J. C **73** (2013) 2480, arXiv: 1304.0790 [hep-ph].

[77] *NLO-NLL Higgsino-like Cross Sections (fully-degenerate case)*, URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections13TeVhino.

[78] S. Agostinelli et al., *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[79] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146 [hep-ph].

[80] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[81] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph].

[82] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

[83] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex].

[84] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **81** (2021) 578, arXiv: 2012.00578 [hep-ex].

[85] ATLAS Collaboration, *Search for supersymmetry in events with four or more charged leptons in 139* fb^{-1} *of* $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, JHEP **07** (2021) 167, arXiv: 2103.11684 [hep-ex].

[86] ATLAS Collaboration, *Searches for new phenomena in events with two leptons, jets, and missing transverse momentum in* 139 fb^{-1} *of* $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, (2022), arXiv: 2204.13072 [hep-ex].

[87] ATLAS Collaboration, *Search for charginos and neutralinos in final states with two boosted hadronically decaying bosons and missing transverse momentum in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **104** (2021) 112010, arXiv: 2108.07586 [hep-ex].

[88] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv: 1510.03823 [hep-ex].

[89] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-023, 2015, URL: https://cds.cern.ch/record/2037700.

[90] ATLAS Collaboration, *Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking*, JHEP **08** (2020) 080, arXiv: 2005.09554 [hep-ex].

[91] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762** (2016) 012036, arXiv: 1608.05806 [physics.data-an].

[92] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv: 1009.1003 [physics.data-an].

[93] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689, arXiv: 2007.02645 [hep-ex].

[94] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 17, arXiv: 1406.0076 [hep-ex].

[95] ATLAS Collaboration, *In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 135, arXiv: 1807.09477 [hep-ex].

[96] *SimpleAnalysis: Truth-level Analysis Framework*, tech. rep., CERN, 2022, URL: https://cds.cern.ch/record/2805991.

[97] CMS Collaboration, *Search for Higgsino pair production in pp collisions at* $\sqrt{s} = 13$ *TeV in final states with large missing transverse momentum and two Higgs bosons decaying via* $H \to b\bar{b}$, Phys. Rev. D **97** (2018) 032007, arXiv: 1709.04896 [hep-ex].

[98] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instrum. Meth. A **434** (1999) 435, arXiv: hep-ex/9902006.

[99] A. L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693.

[100] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, [Erratum: Eur.Phys.J.C 73, 2501 (2013)], arXiv: 1007.1727 [physics.data-an].

[101] L. Heinrich, M. Feickert, G. Stark, and K. Cranmer, *pyhf: pure-Python implementation of HistFactory statistical models*, J. Open Source Softw. **6** (2021) 2823.

[102] L. Heinrich, M. Feickert, and G. Stark, *pyhf: v0.6.3*, version 0.6.3, URL: https://github.com/scikit-hep/pyhf/releases/tag/v0.6.3.

[103] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, (1994).

[104] ATLAS Collaboration, *Search for non-resonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, tech. rep., CERN, 2022, URL: https://cds.cern.ch/record/2811390.

[105] ATLAS Collaboration, *Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector*, tech. rep., CERN, 2021, URL: http://cds.cern.ch/record/2777236.

[106] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV*, Eur. Phys. J. C **75** (2015) 303, arXiv: 1412.7086 [hep-ex].

[107] *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, tech. rep., CERN, 2017, URL: https://cds.cern.ch/record/2261772.

[108] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, tech. rep., CERN, 2019, URL: https://cds.cern.ch/record/2688062.

[109] ATLAS Collaboration, *Combination of searches for non-resonant and resonant Higgs boson pair production in the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}b\bar{b}$ decay channels using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, tech. rep., CERN, 2021, URL: https://cds.cern.ch/record/2786865.

[110] ATLAS Collaboration, *Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2021), arXiv: 2112.11876 [hep-ex].

[111] ATLAS Collaboration, *Constraining the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV*, tech. rep., CERN, 2022, URL: https://cds.cern.ch/record/2816332.

[112] M. Cepeda et al., *Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC*, CERN Yellow Rep. Monogr. **7** (2019) 221, ed. by A. Dainese et al., arXiv: 1902.00134 [hep-ph].